IJCSIS Vol. 8 No. 3, June 2010
ISSN 1947-5500

International Journal of Computer Science & Information Security

Editorial
Message from Managing Editor

The International Journal of Computer Science and Information Security is an English language periodical on research in general computer science and information security which offers prompt publication of important technical research work, whether theoretical, applicable, or related to implementation.

Target Audience: IT academics and business people concerned with computer science and security; university IT faculties; industry IT departments; government departments; the financial industry; the mobile industry and the computing industry.

Coverage includes: security infrastructures, network security: Internet security, content protection, cryptography, steganography and formal methods in information security; multimedia, image processing, software, information systems, intelligent systems, web services, wireless communication, networking and technologies.

Thanks to authors who contributed papers to the June 2010 issue and the reviewers, for providing valuable feedback comments. IJCSIS June 2010 Issue (Vol. 8, No. 3) has an acceptance rate of 30 %.

Available at http://sites.google.com/site/ijcsis/

IJCSIS Vol. 8, No. 3, June 2010 Edition

ISSN 1947-5500

Abstracts Indexed by (among others):

IJCSIS EDITORIAL BOARD

TABLE OF CONTENTS

****

Implementation of SPIN Model Checker for Formal Verification of Distance Vector Routing Protocol

Kashif Javed
Department of Information Technologies
Abo Akademi University
Turku, FIN-20520, Finland
Kashif.javed@abo.fi

Asifa Kashif
Department of Electrical Engineering
National University of Computer and
Emerging Sciences, Islamabad, Pakistan
asifa.ilyas85@gmail.com

Elena Troubitsyna
Department of Information Technologies
Abo Akademi University
Turku, FIN-20520, Finland
Elena.Troubitsyna@abo.fi

***Abstract* - Distributed systems and computing requires routing protocols to meet a wide variety of requirements of a large number of users in heterogeneous networks. DVR is one of many other employed protocols for establishing communication using routes with minimum cost to different destinations from a given source. Research work presented in this paper focuses on implementation of DVR in SPIN and provides formal verification of correctness of DVR behaviour covering all required aspects. Simulation results clearly show a proof of the established paths from each router to different destinations in a network consisting of six routers and a number of links.**

Keywords: Formal Verification, DVR Protocol, SPIN Model Checker, Distance Vector Routing, Implementation in PROMELA

I. INTRODUCTION

A computer network consists of a number of routers which have the capability to communicate with each other. Routing Information Protocol (RIP) is widely used for routing packets from a source to its destination in computer networks. RIP requires information about distance and direction from source to destination. Each router, in the Distance Vector Routing (DVR) methodology, keeps updated record of distances and hops of its neighbours. Various techniques are used to gather useful routing table information for each router. In one approach, special packets are sent by each router and are received back after having time-stamped by the receivers. Chromosomes have been employed in the Genetic Algorithm [1] to select the most optimal path by utilizing its fitness function, selection of next generation and crossover operation for updating the routing tables in an efficient manner. Thus, all routers keep refreshing their routing tables and maintain latest information about other neighbouring routers in order to provide optimized performance in the available network [1-3].

Mahlknecht, Madni and Roetzer [4] has presented an efficient protocol that uses hop count and cost information in its Energy Aware Distance Vector (EADV) routing scheme and makes use of shot-multi-hop routing for consuming lesser energy in the wireless sensor networks. EADV can do well for long lasting battery-powered sensor nodes while using the lowest cost path towards the selected sink node. An algorithm is considered the most effective if it contains the correct and latest information about its neighbours in its DVR table. An effort has been made by Liwen He by devising a computational method to protect a network from internal attacks (such as mis-configuration and compromise) through the use of verifying routing messages in the DVR protocols [5]. Formal verification of standards for DVR protocols has also been comprehensively presented by Bhargavan, Gunter and Obradovic [6] using three case studies. The researchers have used HOL (an interactive theorem prover and SPIN (model checker) to verify and prove salient properties of DVR protocols. HOL and SPIN have been employed by these researchers for providing a proof of convergence for the RIP [7].

The remaining paper is organized as follows. DVR protocol is presented in Section II and Section III describes the use of SPIN tool and PROMELA language for formal verification. System design and implementation has been discussed in Section IV covering network topology, implementation details and operation of DVR protocol. Formal verification of simulation results has been illustrated in Section V and finally conclusions and future work is given in Section VII.

II. DISTANCE VECTOR ROUTING PROTOCOL

A. General Methodology

A routing table is required to be maintained for each router in the network for the purpose of working of a DVR scheme. Routing table information is used to determine the best path (i.e. having minimum cost in terms of distance or hops) from a source to destination. Links are needed to connect concerned routers for establishing communication. An optimal DVR protocol has to exchange frequent messages in order to update the routing table of each router. So, exchanging information among neighbours is carried out on regular intervals.

Routing table of every router keeps necessary information (i.e. id of neighbouring routers, most suitable outgoing link to be used for the destination, distance, hops (number of routers on the route), time delay, number of queued messages on the link). The process of making forwarding decision for selecting the best optimal path from source to destination is based on a combination of these parameters. The objective of routers is to send packets to hosts connected to the networks for heterogeneous requirements of a large number of users. In this way, efficient DVR schemes ultimately establish good global

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

paths by connecting hosts in a distributed environment covering very long distances. Those routers are taken as neighbours which have links/interfaces to a common network.

B. Routing Information Protocol

RIP [8,9] is a widely used protocol for finding the optimal path to the destination in a network. Each router has a routing table and all routers periodically updated their routing tables by using advertising approach. All routes of a router are advertised through the mechanism of broadcasting RIP packets to all the neighbouring routers in the network. Every router checks the advertised information of neighbouring nodes and changes information only in its routing table if the new route to the same destination further improves the existing route length. In other words, the updated routing table information now takes to the best available route so far for the relevant destination.

The number of hops in the RIP are kept low (up to 15) for the route length for faster convergence [6,7]. RIP methodology, however, prevents formation of loops between pairs of routers in order to minimize convergence time as well as permitted route length. Timer expiry record is also maintained in every routing table and is normally set to 180 seconds whenever a routing table is updated. As routers advertise after every 30 seconds, the destination is considered unreachable if a router is not refreshed for 180 seconds. It further waits for another 120 seconds. If the router remains un-refreshed during this time as well, then its route is removed from the routing tables of the concerned routers. This requirement is incorporated to cater for broken links, faulty networks and congestions.

III. Use of SPIN and PROMELA

A. Formal Verification

A number of new systems and methodologies are being devised by the researchers in different areas of science, technology and engineering as a result of meaningful R&D work being undertaken by academic and research institutes all over the world. Every proposed system requires a proof of its correctness by gathering results using simulation and testing techniques. Formal verification terminology [10,11] is in fact a process of actual demonstration of the system in order to check its correctness under the defined boundaries and valid conditions of used parameters/variables.

Precision and accuracy of the system is verified by running the programming modules by employing required algorithms in the model checking approach. Errors occurred (if any) are properly identified under varying conditions so that such errors can be easily located by the users and are later on repaired/tackled by adjusting specifications of the model. Afterwards, the model description is fine tuned to achieve required model specifications for verification of correct results of the system.

B. SPIN Tool and PROMELA High Level Language

SPIN [12,13] is a open-source software tool and is widely used for the formal verification of software systems working in the distributed environment. Inspiring applications of SPIN include the verification of the control algorithms for various applications, logic verification of the call processing software for a commercial data communication, critical algorithms for space missions, operating systems, switching systems, distributed & parallel systems and formal verification of various routing protocols. This tool also supports interactive, random and guided simulations for a wide variety of applications. Spin can be used in four main modes (i.e. as a simulator, as an exhaustive verifier, as a proof approximation system and as a driver for swarm verification).

Spin provides efficient software verification and supports the PROMELA (PROcess MEta LAnguage) high level language to specify systems descriptions [14]. It is a SPIN's input language which is used to build detailed PROMELA models for complete verification of system designs. It provides a way for making abstractions of distributed systems. Different assumptions are used in SPIN to verify each model. After checking correctness of a model with SPIN, it can then be used to build and verify subsequent models of the system so that the fully developed system produces the required behavior. PROMELA programs consist of processes, message channels, and variables.

IV. System Design and Implementation

A. Network Topology

The network topology shown in Figure 1 has been used for implementation of DVR protocol. There are six routers (A, B, C, D, E & F) and seven links (edges). Each link connects two routers. Weight values range from 2 to 23 for different links and these values indicate distances between routers. Integer values have been used and distance units can be chosen during actual implementation of the network. For example, the distance between routers A and C via B is 6 using 2 hops and via D, E and F is 33 using 4 hops.

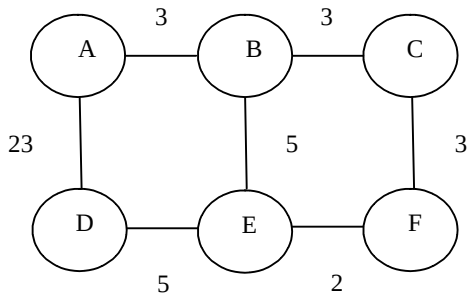


Figure 1: Network Topology

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

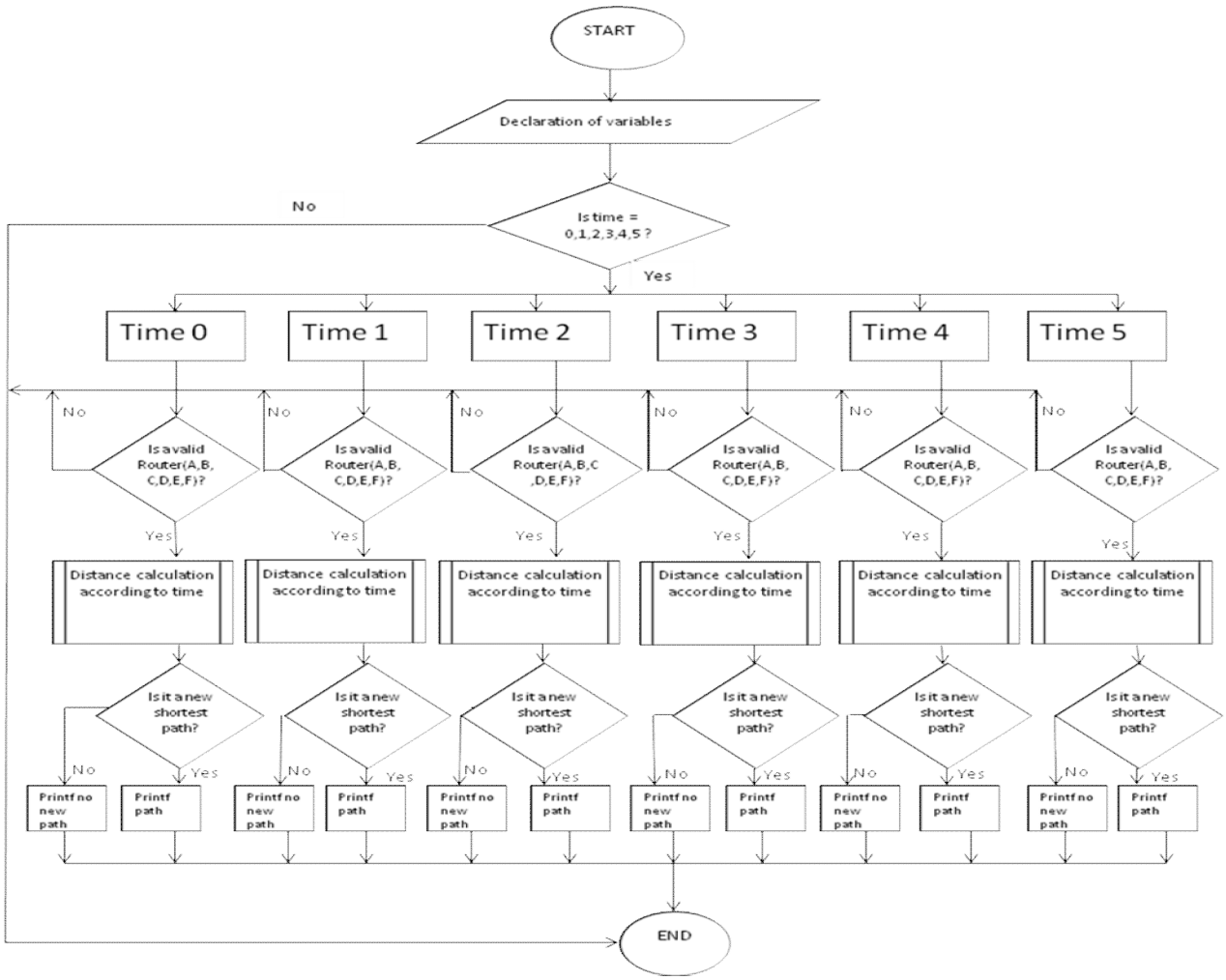


Figure 2: System Flowchart

B. System Implementation

SPIN's PROMELA language has been used to construct complete model of DVR protocol on a Pentium machine. Packets from the source to destination travel using links provided by routers by making use of their routing tables for the given distributed environment of the network. After initialization of the variables, distance is calculated from each router at time period T=0, T=1, T=2, T=3, T=4 and T=5. At each stage it checks whether the measured distance forms a new shortest path or not. Whenever the shortest path is found from the source to destination, routing table entry for the concerned router is automatically updated to make good forwarding decision in order to ensure optimal path, having minimum distance, for faster communication. Thus, each router updates its routing table after each time period. The main objective of the DVR protocol is to provide the current best route (path) from source to destination for each communication. Flowchart of the modeled system in SPIN/PROMELA is shown in Figure 2.

For the given network, the PROMELA program has six processes (one for each time period) to find distance based upon the time period conditions (0 to 5). The found distance from a particular source to destination for each time period is compared with all the available alternate routes. Router's table is only updated if the new distance is minimum between the selected source and destination. The new shortest path is recorded after each calculation. If the determined route does not find minimum distance during the given time period, then it ignores its path without updating any entry in the routing table. Routers improve their routes whenever a router advertises its routing table to its neighbours. So, new routes are determined purely based on their length measured in distance. For timely convergence, the number of hops involved in the length are limited to 15 as already highlighted by Bhargavan et. al. [7].

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

		Via							Via							Via					
	From A	A	B	C	D	E	F	From B	A	B	C	D	E	F	From C	A	B	C	D	E	F
T=0	A							A	3						A						
	B		3					B							B		3				
	C							C			3				C						
	D				23			D							D						
	E							E					5		E						
	F							F							F						3
T=1	A							A							A		6				
	B							B							B						
	C		6					C							C						
	D							D	26				10		D						
	E		8		28			E							E		8				5
	F							F			6		7		F						
T=2	A							A					33		A						
	B				33			B							B						10
	C							C					10		C						
	D		13					D							D		13				10
	E							E	31		8				E						
	F		9		30			F							F		10				
T=3	A							A							A		36				13
	B							B							B						
	C		13		33			C							C						
	D							D			13				D						
	E		11					E							E		34				
	F							F	33						F						
T=4	A							A			36				A						
	B				36			B							B						36
	C							C	36						C						
	D		16					D							D						36
	E							E							E						
	F				39			F							F		36				
T=5	A							A							A						
	B							B							B						
	C							C							C						
	D							D							D						
	E							E							E						
	F							F							F						

Table1: Calculated Distance from Routers A, B and C for Different Destinations at Time Periods T=0 to T=5

C. Operation of DVR Protocol

DVR protocol works independently for every destination and it is assumed that there is no topology change for protocol's convergence during every time period. The router broadcasts after every 30 seconds and the destination is taken as inaccessible if it is not refreshed for 180 seconds. The route is removed from the tables of concerned routers if the particular router fails to refresh itself for 300 seconds.

Although the PROMELA's built model can be used for any number of routers but its operation is restricted only to the topology given in Figure 1. For the purpose of explanation of the model, it is assumed that every router operates without any problem and updates its routing table during regular intervals of time.

At Time=0, it calculates distances to neighbouring routers from each router having maximum one hop. Thus, distance from A to B is 3 & A to D is 23from router A; from router B it is 3, 3 & 5 for routers A, C & E respectively; and distances are 5, 5 & 2 for routers B, D & F respectively from router E. These distances can be observed in Tables 1 and 2. Now two hops from the current router are taken for T=1. So, distance from A

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

		Via							Via							Via					
	From D	A	B	C	D	E	F	From E	A	B	C	D	E	F	From F	A	B	C	D	E	F
T=0	A	23						A							A						
	B							B		5					B						
	C							C							C			3			
	D							D				5			D						
	E					5		E							E					2	
	F							F						2	F						
T=1	A							A		8		28			A						
	B	26				10		B							B			6		7	
	C							C		8				5	C						
	D							D							D					7	
	E							E							E						
	F					7		F							F						
T=2	A					13		A							A			9		10	
	B							B				31		8	B						
	C	29				10		C							C					10	
	D							D		31					D						
	E	31						E							E			11			
	F							F		11					F						
T=3	A							A						11	A						
	B					13		B							B					33	
	C							C				34			C						
	D							D							D			16		33	
	E							E							E						
	F	32				16		F							F						
T=4	A					16		A							A			39			
	B							B							B						
	C	36						C							C					36	
	D							D						34	D						
	E	34						E							E			37			
	F							F				37			F						
T=5	A							A							A						
	B							B							B						
	C							C							C						
	D							D							D						
	E							E							E						
	F							F							F						

Table 2: Calculated Distance from Routers D, E and F for Different Destinations at Time Periods T=0 to T=5

to C via B is 6; A to E via D 28; and A to E via B is 8 as given in Table 1.

When T is taken as T=2, three hop lengths are counted for determining the distance from each router. From router D, measured distances are 13 via E to A, 29 via A to C, 10 via E to C and 31 via A to E. Same can be seen in Table 2. Hop length is four when T=3, distance covered to B via E, D via C and D via E is 33, 16 and 33 respectively from router F as shown in Table 2. Similarly, routes have distances of 36 (B via F), 36 (D via F) and 36 (F via B) from router C (for T=4) as given in Table 1. Both Tables 1 and 2 clearly indicate that no routes are available from any router when T=5 (six hops) for network configuration of Figure 1.

V. Formal Verification of Simulation Results

The implemented system in PROMELA programming language has been tested exhaustively and obtained simulation results are shown in Tables 1 and 2. Spin model checker has been used to verify all the results. The developed model ensures that all the routers correctly maintain and update their tables as and when new routes are searched and visited. The broadcast mechanism works well at different time periods and the system provides correct and optimized results from each router to various destinations depending upon network topology, layout of routers and links connecting different routers in the network.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

The SPIN's verification model successfully checks all the available routes via different routers and permits only the shortest path from the available options. It is evident from the following decisions (only four out of many are presented here):

1) *At T =1, the route length from E to C via B is 8 where as it is 5 via F. So, E router adopts F router's path to reach C.*
2) *The distance between routers B & E via A and via C is 31 and 8 respectively. SPIN's checker confirms that minimum distance is covered for reaching to C from E when T=2.*
3) *When T=3, the path cost determined by the model is 13 from C to A via F, E & B but another path for connecting the same two router via B, E & D is 36, each path makes use of four hops. Of course, the longer path is simply ignored.*
4) *Similarly, route length from F to D through C, B & A is 32 and it is 16 via routers C, B & E. A saving of 16 is noted while using the most economical path.*

A careful analysis of the simulation results shown in Tables 1 & 2 clearly indicates that the modeled system in PROMELA operates correctly and provides the best possible routes involving minimum distances using DVR protocol on the given network environment. The system works efficiently under all conditions and the SPIN model checker has guaranteed correctness of all results. It means that all the routing tables are timely updated while messages are being sent to various destinations from a particular source. Now, this can be extended to bigger networks in the distributed environment for efficient and correct functioning using SPIN tool.

VI. Conclusions and Future Work

Many researchers have implemented DVR protocols for various applications. In this research work, PROMELA language has been used to implement DVR protocol on a six router model. Formal verification of DVR protocol properties has been shown through the use of SPIN checker model. The simulation results amply demonstrate correctness and reliability of DVR protocol under varying conditions. Performance of the implemented has been extremely well and it can further be improved to make it more efficient in terms of reducing storage space requirements, incorporating security mechanism for safer communication, minimizing congestion at peak loads and making it fault-tolerant for enhancing its reliability and flexibility.

References

[1] M. R. Masillamani, A. V. Suriyakumar, R. Ponnurangam and and G.V.Uma, "Genetic Algorithm for Distance Vector Routing technique", AIML International Conference, 13-15 June 2206, Egypt, pp. 160-163.

[2] Andrew S.Tanenbaum, "Computer Networks", 4th Edition,. Prentice-Hall Inc., 2005.

[3] G. Coulouris, J. Dollimore and T. Kindberg, "Distributed Systems : Concepts and Design, 4th Edition, Addison-Wesley, 2005.

[4] S. Mahlknecht, S. Madani and M. Rötzer, "Energy Aware Distance Vector Routing Scheme for Data Centric Low Power Wireless Sensor Networks," *Proceedings of the IEEE International Conference on Industrial Informatics INDIN 06*, Singapore, 2006.

[5] Liwen He, "A Verified Distance Vector Routing Protocol for Protection of Internet Protocol", Lecture Notes in Computer Science, Networking – ICN 2005, Volume 3421, Springer, pp. 463-470.

[6] K. Bhargavan, D. Obradovic and C. A. Gunter, "Formal Verification of Standards for Distance Vector Routing Protocols", Journal of the ACM, Vol. 49, no. 4, July 2002, pp. 538-576.

[7] K. Bhargavan, C. A. Gunter, and D. Obradovic, "Routing Information Protocol in HOL/SPIN", Proceedings of the 13th International Conference on Theorem Proving in Higher Order Logics 2000, August 14 - 18, 2000, London, UK, pp. 53-72.

[8] C. Hendrick, "Routing Information Protocol", RFC 1058, IETF, June 1988.

[9] G. Malkin, 'RIP Version Carrying Additional Information', IETF RFC 1388, January 1993.

[10] J. Katoen, "Concepts, Algorithms and Tools for Model Checking", Lecture Notes 1998/1999, Chapter1: System Validation.

[11] N. A. S. A. Larc, "What is Formal Methods?", http://shemesh.larc.nasa.gov/fm/fm-what.html, formal methods program.

[12] R. de Renesse and A. H. Aghvami "Formal Verification of Ad-Hoc Routing Protocols using SPIN Model Checker", Proceedings of IEEE MELECON'04, Croatia, May 2004.

[13] G. J. Holzmann, "The Model Checker SPIN", IEEE Transactions on Software Engineering, Vol. 23, No. 5, May 1997, pp. 279-295.

[14] G. J. Holzmann, "Design and Validation of Computer Protocols", Prentice Hall, November 1990.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Integrated Queuing based Energy-Aware Computing in MANET

Dr. P.K. Suri, Dean, Faculty of Sciences, Dean, Faculty of Engineering, Professor, Deptt. of Comp. Sci. & Appl., Kurukshetra University
Kurukshetra, Haryana, India
pksurikuk@rediffmail.com

Kavita Taneja, Assistant Professor
M. M. Inst. of Computer Tech. & Business Mgmt.
Maharishi Markandeshwar University
Mullana, Haryana, India
kavitatane@gmail.com

***Abstract*— Mobile Computing has witnessed a flare-up of applications in mobile and personal communication. It is a phrase that embodies on-the-go business initiatives. By marrying today's dominant office computing environment with increasingly compact-but-powerful handheld devices, mobile computing makes it possible for millions of workers to conduct business on their feet and from the road. Energy efficiency is an important design consideration due to the limited battery life of mobile devices. In order to minimize energy consumption and maximize the network life time, the proposed simulator opt for intelligent routing through substitute routes instead of conventional routing through shortest route. To reduce energy consumption only devices in the traversed route are active, other mobile devices of network are switched off. Also simulator implements clusters for efficient sleep and active mobile device mechanism and reflects on MANET in terms of queuing network and consider the packets arrival rate in terms of poisson distribution.**

Keywords- Queuing theory, poisson distribution, clustering, mobile unit, mobile ad hoc network.

I. INTRODUCTION

MANET (Mobile Ad Hoc Network) is a group of mobile units (MUs) that instantly form a network among themselves without any fixed infrastructure. The mobile computing offers users in such open networks, the on the fly access to information. Due to the explosion of wireless and portable devices, such as cellular phones, PDAs (Personal Digital Assistants), palm computers, user is spoilt to enjoy the freedom and convenience in their daily lives [1]. In addition, the race against time and enhancements in the current MANETs make user more expected to do things through the Internet, such as online banking, online shopping etc. Then there is no doubt user's demand for information access in mobile environment increases spectacularly [2]. Since the mobile computing is relatively novel to its big brother– desktop computing, the resource in the mobile computing world is relatively restricted. Also constraints, such as limited hardware resources especially bounded battery life, stochastic topology hence uncertain device location, etc, make challenges for optimally utilizing resources. Increasingly, power consumption within such open networks is becoming a core issue for the low-power mobile devices [3]. For example, soldiers in a battlefield can have handheld mobile devices to communicate with each other, or in emergency situations like earthquakes where the existing infrastructure has been destroyed, an ad hoc network can instantly be deployed to aid in disaster recovery, meetings or conference in which persons wish to quickly share information, and data acquisition operations in inhospitable terrains. In order to crack shortages in mobile computing for smooth satisfaction of the emerging needs, the enhanced computing needs to be adaptive [4]. The primary goal of routing protocol in MANET is correct and efficient route establishment between a pair of MUs so that packets can be delivered in a timely manner with minimum battery consumption for longer network connectivity. In the evaluation of any algorithm for energy conservation, an estimate of energy consumption is necessary [5].

The energy is consumed per MU in two phase as shown in Eq. (1), the energy used in route discovery and amount of energy used in packet transmission.

$$E^{\text{Total}} = E^{\text{Route-Discovery}} + E^{\text{Packet-Transmission}} \quad (1)$$

During packet transmission a typical MU may exist in any one of the four modes:

a) *Sleep mode*: Sleep mode has very low power consumption. The network interface at a MU in sleep mode can neither transmit nor receive packets; the network interface must be woken up to idle mode first by an explicit instruction from the MU.
b) *Idle mode*: In this mode MU is neither transmitting nor receiving a packet. This mode still consumes least amount of power because the MU has to listen to the wireless medium continuously in order to detect a packet that it should receive, so that the MU can then switch into receive mode.
c) *Receive mode*: In this mode a MU actually receives a packet.
d) *Transmit mode*: In transmit mode a MU uses its maximum power level to transfer packet to another MU.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Receive and idle mode require similar power, and transmit mode requires slightly greater power [6]. Sleep mode requires more than an order of magnitude less power than idle mode. These measurements show that the network interface expends similar energy, whether it is just listening or actually receiving data. Hence, switching to sleep mode whenever possible is a wise decision on the part of MU that will lead to momentous energy savings. The sources of power consumption, with regard to network operations, can be classified into two types: (a) communication related, and (b) computing related. Communication involves the cost of sending data along with control packet from source MU, routed through intermediate MU and cost at receiving at destination end. It mainly deals with the cost of routing in network or in route establishment. Whereas computing mainly involves the cost of using CPU and main memory, disk drives and other components of computer for calculating computing related cost [7]. If we take the shortest route to deliver the message it is not the good way of energy conservation in network. As a group of MUs coming under the shortest route is used rapidly as compared to other MUs, so their power decreases rapidly and may a situation come that they are having no power. In that situation we refer the substitute path routing [8]. The substitute routes and clustering approach ensures that the energy of a group of MUs is used at a particular instant in a respective way [9]. By this all MUs of a network take active participation in route selection; overall energy expenditure of a network is minimized, causing maximization of network life. The remainder of this paper is organized as follows. The basic concept of queuing theory and MANETs is explained in section II. Section III gives the Simulator for finding substitute route and clustering approach in MANET. Section IV gives the probabilistic approach of finding the number of sleep MU and active MU. Finally we conclude our paper in section V.

II. QUEUING THEORY AND MANET

We consider that inter arrival times of packets at a MU follow a discrete poisson distribution in which an event occurring exactly k times during an interval t is given by a probability mass function

$$g_k(t) = (t\lambda)^k \ (1/k!)\ e^{-t\lambda}. \quad (2)$$

where λ is the average number of times the event occurs in a unit period. In our network the packet is always transferred from source to destination, we consider our source as a server. With single server, service times follow negative exponential distribution. Packets are served in FIFO order. Pseudo random numbers are generated and using χ^2 test, we have samples (T) from an exponential distribution with specified expected value 1/λ as

$$RN=RNDY1(DUM), \quad (3)$$
$$T = -1.0/\lambda*ALOG(RN). \quad (4)$$

We generate T_i's according to Eq. 4 and keep on adding them till their sum exceeds 1 and the count gives the Poisson sample (k). At any given time t, the probability of a MU busy in packet processing is given by "(5)".

$$\rho = \beta / \alpha. \quad (5)$$

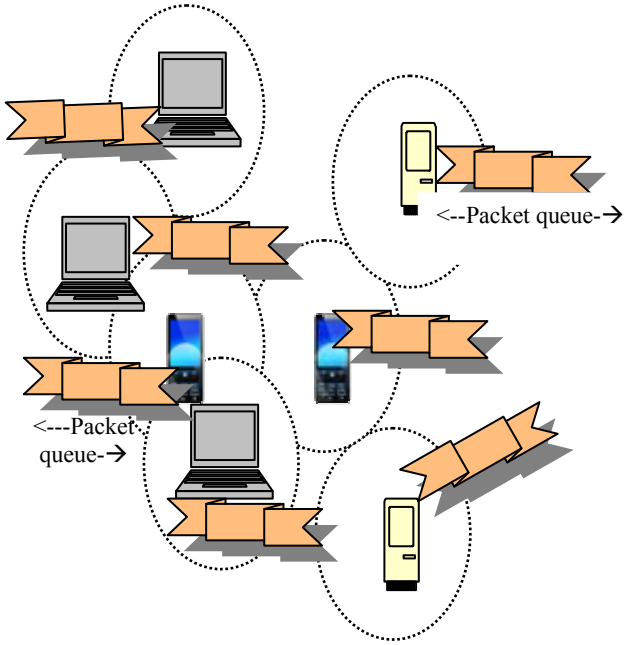


Figure1 Queuing theory embedded in MANET

Considering computation of Packet queue length of each MU as shown in figure 1, if more packets passed through Packet queue it is lost, here Packet queue limit is n. Because packet arrival at a MU is highly stochastic, hence simulator design assumes that the Packet queue length is such that it is 99% of the time sufficient to hold the packets waiting to be processed. Thus the probability of having more than n packets in wait for MU is

$$\rho^{n+1}=0.01. \quad (6)$$

In "(6)", n gives the long enough Packet queue length with 99 per cent assurance. For efficient routing in order to lessen the number of packets loss we suggest that the packet to be delivered in substitute routes instead of shortest route. So initially the numbers of all possible substitute routes are to be calculated. Then while packets traveling through a particular route the MUs of other routes to be switched off if they are not doing any useful work in order to minimize energy consumption. The algorithm describing this substitute route finding procedure is given in section III.

III. ALGORITHM FOR SUBSTITUTE ROUTE AND CLUSTERING APPROACH.

In this section we propose our algorithm to find out substitute routes in a network to maximize network life if direct route between source and destination does not exist [10]. Initially the MUs are set up in an area to establish network connectivity. Then we go for substitute route finding

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

procedure, by running this algorithm in a network we keep track of all substitute routes which does not contain any duplicate MUs. By taking the substitute routes traffic load is shared and congested routes are avoided which may cause in

Algorithm for substitute route ( ):

1. Initialize all MUs of network to READY state from SWITCH OFF state at time of deployment of network.
2. Put the source MU (S) in Packet queue.
3. Repeat step 4 until Packet queue is empty.
4. Process front MUs of Packet queue by adding its neighbor's to Packet queue where each entry is unique along with that keep track of their parent MU.
5. End of step 3.
6. As destination (D) is reached stop and find the route traversing from 'D' in a reverse order by tracking the origin till 'S' at origin is reached.
7. SWITCH OFF all MUs coming in the route for finding another route without any duplicate MU.
7. Go to step 4
8. Exit.

Figure 2. Algorithm: Substitute Route

retransmission due to collision. Since no MU is duplicated, all MUs take in active participation in route selection. No MU is penalized more as compare to other MUs in network. So there is no network partition which causes maximization of network life [4, 11, 12]. Before going to the algorithm find out list of neighbors of all MUs. For evaluation of algorithm a priority Packet queue is considered. Considering that the front MU of a Packet queue has to be processed first as it has highest priority than other. The network is divided level wise. Each level is forming a cluster [6, 13]. Each cluster is assigned with a cluster head (CH). A MU is selected as a CH having more battery energy [14] which can communicate to its neighbor in one hop distance. The lower level of network which is nearer to the source MU keeps track of all substitute routes. All the CHs which are formed in a network can exchange their messages in between them at any point of time by broadcasting messages in between the CHs. The CH in each cluster broadcasts the control message to its entire MUs. This control message contains the information about MU ID, and residual energy. When packet starts transferring in one route, the CH of initial level make the MU taking part in route selection as ACTIVE and make other MUs to be in SLEEP mode. It broadcasts the message to all the CHs. All the CH makes the MUs in that path active and other to be in sleep mode. When the next packet starts traversing it transfer through another route. Then the CH positively play their role to make the MU falling in that route to ACTIVE and other MUs of the cluster to be SLEEP mode. Using this when packets transferring in a particular route, the MUs of that route are in ACTIVE state where all other MUs of a network are in SLEEP mode. So this lessens the overall energy consumption of whole network. At any point of time if CH is not having required amount of energy it can elect the MU with huge resource as CH in order to maintain network connectivity.

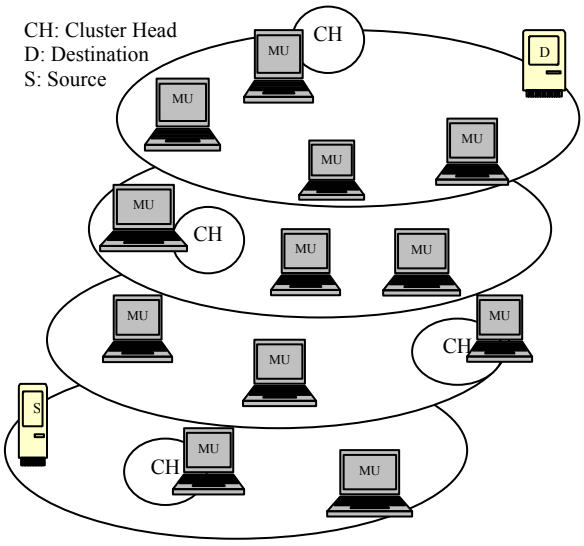


Figure3 Clustering in MANET

IV. THE PROBABILISTIC APPROACH

To find out the number of MUs to be ACTIVE and the number of MUs to be in SLEEP MU, we take the following probabilistic approach. Let T be the total number of MUs, A be the number of ACTIVE MUs and S be the number of MUs in SLEEP mode. Now we find the joint probability of 'a' MUs among the total MUs active and 's' MUs among total MUs in sleep mode. The packet arrival distribution follows Poisson's distribution (λ). Let probability of a MU being ACTIVE mode= P. The probability of a MU being in SLEEP mode= 1-P. For the discrete random variable, the joint probability mass function

$$P\{A=a \text{ and } S=s\} = P\{S=s \mid A=a\} * P\{A=a\}$$
$$= P\{A=a \mid S=s\} * P\{S=s\}. \quad (6)$$

Also, $\Sigma \Sigma P\{A=a \text{ and } S=s\}=1$, Since A and S are independent

$$P\{A=a \text{ and } S=s\} = P\{A=a\} * P\{S=s\}$$
$$= (\lambda^a e^{-\lambda}/a!) * (\lambda^s e^{-\lambda}/s!). \quad (7)$$

It shows that probability of 'a' number of MUs in ACTIVE state and 's' number of MUs in SLEEP state. By putting the value of probability to 1 and specifying the rate of arrival λ we find the ratio of number of ACTIVE MUs and SLEEP MUs, which effectively measure the energy consumption rate. As by the MU going into SLEEP state it consume less energy, therefore as the number of MUs not doing any useful work in the system going into SLEEP state, then the energy conservation will increase, which enhance our system performance with maximized network lifetime.

V. CONCLUSION

In this paper we give a brief study about role of queuing theory in MANET. As all MUs are battery operated, and battery power is restricted we go for proficient use of this in order to minimize energy consumption and maximize network

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

life [15]. The today's user rely on mobile computing for a multitude of operations as is equipped with range of devices like notebook computers, personal digital assistants and other communication hungry portable stuff, even kitchen of homemakers is gripped with spectrum of wireless interfaces for networked communication [16]. By using the substitute route and clustering technique we try minimize energy consumption of overall network. Also switching some of MUs to SLEEP mode we try to further minimize energy consumption in network. To calculate the number of ACTIVE MUs and SLEEP MUs in network we are using the probabilistic approach. This model promises to provide stochastic handling equipped design platform for extension to an energy efficient protocol for MANET.

REFERENCES

[1] Sarkar, S. K., T.G. Basavaraju, and C. Puttamadappa, "Ad Hoc Mobile Wireless Networks: Principles, Protocols and Applications", Auerbach Publications, 2007.

[2] Do-hyeon Lee , Song Nan Bai , Jae-il Jung, "Enhanced next hop selection scheme for QoS support in multi-hop wireless networks", Proceedings in the 2009 International Conference on Hybrid Information Technology, pp.587-593, August 27-29, 2009, Daejeon, Korea.

[3] D. Kim, J. Garcia and K. Obraczka, "Routing Mechanisms for Mobile Ad Hoc Networks based on the Energy Drain Rate", IEEE Transactions on Mobile Computing. Vol 2, no 2, 2003, pp.161-173.

[4] Thomas Kunz, "Energy-efficient routing in mobile ad hoc networks: a cautionary tale", International Journal of Autonomous and Adaptive Communications Systems, Vol.2 no.1, pp.70-86, March 2009.

[5] L. Feeney and M. Nilsson, "Investigating the Energy Consumption of a Wireless Network Interface in an Ad Hoc Networking Environment", Proceedings in INFOCOM 2001, Vol. 3, pp. 1548–1557, Anchorage, Alaska, April 2001.

[6] C. E. Jones, K. M. Sivalingam, P. Agrawal and J. C. Chen, "A survey of Energy Efficient Network Protocols for Wireless Networks", Journals in Wireless Networks, Vol. 7, no. 4, pp. 343-358, August 2001.

[7] M. Tarique, K. E. Tepe and M. Naserian, "Energy Saving Dynamic Source Routing for Ad Hoc Wireless Networks", Proceedings in Third International symposium on Modelling and Optimization in Mobile Ad Hoc, and Wireless Networks, Canada, pp. 305-310, April 2005.

[8] C.K. Toh. "Maximum Battery Life Routing to Support Ubiquitous Mobile Computing in Wireless Ad Hoc Networks", IEEE Communications Magazine, Vol 39, no 6, 2001, pp.138-147.

[9] E. M. Royer and C. K. Toh, "A Review of Current Routing Protocol for Ad Hoc Mobile Wireless Networks", Proceedings in IEEE Personal Communication Magazine, Vol. 6, no. 2, pp. 46-55, April 1999.

[10] M. Tarique, K. E. Tepe and M. Naserian, "Energy Saving Dynamic Source Routing for Ad Hoc Wireless Networks", Proceedings in Third International symposium on Modelling and Optimization in Mobile Ad Hoc, and Wireless Networks, Canada, pp. 305-310, April 2005.

[11] K. Pappa, A. Athanasopoulos, E. Topalis and S. Koubias, "Implementation of power aware features in AODV for ad hoc sensor networks. A simulation study", Proceedings in IEEE conference on Emerging Technologies and Factory Automation, Rome, pp. 1372-1375, September 2007.

[12] C. Jie, C. Jiapin and L. Zhenbo, "Energy-efficient AODV for Low Mobility Ad Hoc Networks", Proceedings in IEEE International Conference on Wireless Communications, Networking and Mobile Computing, Shanghai, pp. 1512-1515, September 2007.

[13] S. Soro and W. B. Heinzelman, "Cluster head election techniques for coverage preservation in wireless sensor networks", Proceedings in Elsevier Journal Ad Hoc Networks, Vol. 7, no. 5, pp. 955-972, July 2009.

[14] X. Hou, D. Tipper and S. Wu, "A Gossip-Based Energy Conservation Protocol of Wireless Ad Hoc and Sensor Networks", Journals of Networks and Systems Management, Computer Science, Springer New York, Vol. 14, no. 381-414, September 2006.

[15] Zhao Cheng, Wendi B. Heinzelman, "Discovering long lifetime routes in mobile ad hoc networks", Ad Hoc Networks, Vol.6 no.5, pp.661-674, July 2008.

[16] Helmut Hlavacs , Karin A. Hummel, Roman Weidlich , Amine M. Houyou, Hermann De Meer, "Modelling energy efficiency in distributed home environments", International Journal of Communication Networks and Distributed Systems, Vol.4 no.2, pp.161-182, January 2010.

Dr. P. K. Suri received his Ph.D. degree from Faculty of Engineering, Kurukshetra University, Kurukshetra, India and Master's degree from Indian Institute of Technology, Roorkee (formerly known as Roorkee University), India. Presently He is Dean, Faculty of Sciences, Dean, Faculty of Engineering, Kurukshetra University and is working as Professor in the Department of Computer Science & Applications, Kurukshetra University, Kurukshetra, India since Oct. 1993. He has earlier worked as Reader, Computer Sc. & Applications, at Bhopal University, Bhopal from 1985-90. He has supervised six Ph.D.'s in Computer Science and thirteen students are working under his supervision. He has more than 110 publications in International / National Journals and Conferences. He is recipient of 'THE GEORGE OOMAN MEMORIAL PRIZE' for the year 1991-92 and a RESEARCH AWARD –"The Certificate of Merit – 2000" for the paper entitled ESMD – An Expert System for Medical Diagnosis from INSTITUTION OF ENGINEERS, INDIA. His teaching and research activities include Simulation and Modeling, SQA, Software Reliability, Software testing & Software Engineering processes, Temporal Databases, Ad hoc Networks, Grid Computing and Biomechanics.

Kavita Taneja has obtained M.Phil(CS) from Alagappa University, Tamil Nadu and Master of Computer Applications from Kurukshetra University, Kurukshetra, Haryana , India. Presently, she is working as Assistant Professor in M.C.A. at M.M.I.C.T & B.M., M.M. University, Mullana, Haryana, India. She is pursuing Ph.D in Computer Science from Kurukshetra University.

She has published and presented over 10 papers in National/International journals/conferences and has bagged BEST PAPER AWARD, 2007 at International Conference for the paper entitled "Dynamic Traffic -Conscious Routing for MANETs" at DIT, Dehradun. She has supervised five M.Phil scholars in Computer Science. Her teaching and research activities include Simulation and Modeling and Mobile Ad hoc Networks.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

A Review of Negotiation Agents in e-commerce

Sahar Ebadi*
Department of Information System, Faculty of Computer Science and Information Technology, University Putra Malaysia, Serdang, Malaysia
Sah_ebadi@yahoo.com

Md. Nasir Sulaiman
Department of Information System, Faculty of Computer Science and Information Technology, University Putra Malaysia, Serdang, Malaysia
nasir@fsktm.upm.edu.my

Masrah Azrifah Azmi Murad
Department of Information System, Faculty of Computer Science and Information Technology, University Putra Malaysia, Serdang, Malaysia
masrah.azrifah@gmail.com

***Abstract*—With the growth of World Wide Web and the increasing human demand on online trading, there is a pressing need for complex systems that are capable of handling the client needs in e-commerce. In recent years, numbers of Multi Agent System (MAS) developers arise to fulfill this mission by performing a huge number of studies on agent negotiation systems. However, far too little attention has been paid to provide a rich review as a repository for developers to distinguish the aspect and scope of MAS. The purpose of this paper is to do a review of progressing agent negotiation technology in e-commerce. In order to achieve our aim we propose different classification schemata and interpret different models according to the proposed classifications. Popular methods for optimizing negotiation agents have been introduced and the effect of relative techniques has been analyzed. The result of analysis shows that genetic algorithm is the most effective learning technique in optimizing negotiation models. Moreover, we interpret the most prominent negotiation models according to the main parameters on which any negotiation agent model depends. The result of these analysis supplies a resource of differentiating competing alternatives for the area of negotiation agent's models to exploit. Finally, a range of open issues and future challenges are highlighted.**

Keywords-component; Artificial Intelligence; Agent; Multi-Agent System; Negotiation; e-Commerce

I. Introduction

With the rapid growth of the World Wide Web and huge demand of online trading every day, there is a need for complex systems that are capable of addressing the online trading needs of human. Such systems must be capable of establishing communications, making decisions and handling customer's requirements. Many researchers in Multi Agent System (MAS) and agent negotiation systems succeed to fulfill these obligations on e-commerce.

As the number of research conducted on agent negotiation development has rapidly increased, the need of conducting a comprehensive review on negotiation agent in e-commerce has increased. So far, however, there has been little review about agent negotiation models specifically in e-commerce. The purpose of this paper is to review studies conducted on negotiation agent systems in e-commerce (online trading). In general, lack of universally accepted definitions in negotiation agent systems is one of the difficulties in this area. Therefore, we initially clarify the negotiation agent system's characteristics and concepts used in this field. Then we explicate the most popular methods in optimizing negotiation agent models. Related prominent techniques will be explained and some exemplar stand out models will be interpreted.

This paper presents two novel classifications, namely classification according to the type of agents, and classification according to the attitude of agents; in addition to the Wooldridge's popular classification according to the number of agents involved in negotiation. These classifications help developers and researchers to have a better understanding of aspects and scope of agent negotiation systems. These features can facilitate future developments on negotiation mechanism in MAS. Finally, we will trace the progress of negotiation systems generations by analyzing the most prominent works during the last decade, from early 1996 to late 2009. In this work, we will interpret the whole negotiation agent model according to four important characteristics of the system so-called negotiation protocol, negotiation strategy, agent characteristics and negotiation setting. In addition, the most effective learning technique will be verified through interpreting the final utility of exemplar models. The remainder of this article is structured as follows. Section 2 presents essential characteristics and concepts in the scope of our work. Section 3 presents popular methods of optimizing negotiation models and a brief description of the relative techniques. In section 4, proposed classifications are discussed. Section 5 draws together the two later strands. It discusses some exemplar negotiation agent-based applications and highlights new direction to the future works. Section 6 presents the conclusion of this paper.

II. Concepts and Characteristics

As mentioned earlier, one of the difficulties in agent negotiation systems is the lack of universally accepted definitions. In order to draw a clear picture of concepts and characteristics of negotiation agents, it is important to recap some key concepts and definitions which are accepted by some experts in this field.

Agent: an *agent* is a computer system that is situated in some environment, and is capable of autonomous action in this environment in order to meet its designed objective [1]. According to Wooldridge and Jennings [1], the term agent is most generally used to denote a hardware or (more usually)

* *Responsible author*

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

software-based computer system that enjoys the following properties:

- *autonomy*: agents operate without the direct intervention of humans or others, and have some kind of control over their actions and internal state [2]
- *social ability*: agents interact with other agents (and possibly humans) via some kind of *agent-communication language* [3]
- *reactivity*: agents perceive their environment, (which may be the physical world, a user via a graphical user interface, a collection of other agents, the INTERNET, or all of these combined), and respond in a timely fashion to changes
- *pro-activness*: agents do not simply act in response to their environment, they are able to exhibit goal-directed behavior by *taking the initiative*.

In most of the real world problems, agents need to interact with other agents to achieve their objectives. Many problem cases are innately multi party or social such as negotiation scenarios. Many are more complex to be solved by an agent (e.g. monitoring and performing an electronic marketplace). In these cases, multi agent systems are designed to address these issues.

Multi Agent System: generally Multi Agent Systems (MAS) refers to such a system that many intelligent agents involved interact with each other in a selfish or cooperative manner. In MAS, agents are autonomous entities and can cooperate to reach a common goal (e.g. an ant colony) or just follow their own preferences (as in the free market economy).

According to Sycara [4], "the characteristics of MASs are that (1) each agent has incomplete information or capabilities for solving the problem and, thus, has a limited viewpoint; (2) there is no system global control; (3) data are decentralized; and (4) computation is asynchronous."

Since trading domains are often bounded by limited resources and abilities, negotiation has become an essential activity in e-commerce applications.

Negotiation: negotiation is a process in which two or more parties with different criteria, constraints, and preferences, jointly reach to an agreement on the terms of a transaction [5].

Negotiation Agents : according to Raymond [6], the notion of agency can be applied to build robust, fast and optimal architecture for automated negotiation systems within a group of software agents communicating and autonomously making decisions on behalf of their human users .

After a clear understanding of what agent-based concepts are, it is necessary to emphasize the key concepts on optimizing negotiation agent systems on e-commerce.

III. NEGOTIATION AGENT'S CLASSIFICATION

Negotiation Agents can be categorized by severed orthogonal dimensions. In this work, we present three classifications. The first one is a popular classification introduced by Wooldridge [5]. The second and third classifications are proposed by this work. Proposed classifications schemata are done according to i) the types of agents involved in negotiation and ii) the attitudes of agents involved in negotiation.

A. Number of Agents Involved in the Negotiation

This category is one of the most clear and popular categories which are divided into three groups named one-to-one negotiation, one-to-many negotiation and many-to-many negotiation.

One-to-one negotiation is suitable in situations where one agent is negotiating with another agent (e.g. Kasbah model proposed by Anthony et al. [7] and Ryszard Kowalczyk [8]) in case where one agent is involved in the negotiating with other agent. This is a simple but basic kind of negotiation on e-commerce.

One-to-many negotiation is when one agent negotiates with several agents as its opponents. This kind of negotiation, in fact, is originated from several combinatorial one-to-one negotiations. A practical example of a one-to-many negotiation is auction system where several bidders participate in an auction at the same time. Many researchers [9, 10] on online trading propose their models based on this classification.

Many-to-many negotiation as described by Wooldridge [5], is when many agents negotiate with many other agents simultaneously. Jiangbo many-to-many negotiation framework is one of the best examples of this category [11-13].

B. Type of Agents Involved in Negotiation

We believe finding the suitable classification of opponent agents type will help agents to choose the best possible tactics to deal with their opponents over a particular good. We think this classification will decrease the negotiation cost by improving the accuracy of agent's beliefs. Many researchers [10, 14, 15] have discussed types of agents involved in a negotiation. According to the variety of application different concepts are offered. Nguyen [10] divided agents involved in negotiation into two groups according to the amount of the concede which they are willing to give. The first type is conceder and the second is non-conceder. The conceder is referred to the agents that are willing to concede with the aim of selling. In contrast, non-conceder agents are those who just deal with a situation where there is some amount of benefit. Bayesian classification was employed to identify the probable types of agents. Then, agents try to modify appropriate strategies according to the type of the opponent agents.

A new direction in this area is to apply some sense of machine learning techniques to increase accuracy of the opponent agent classification. Therefore, agent is able of making a better deal that improves the performance of the model.

Ros and Sierra [16] defined 5 types of agents that have been experienced by the model and finally evaluated by utility

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

product and utility difference. These 5 types of agents so-called NegoTo agent, Random agent, Alternate agent, TOagent and Nego agent are differentiated by tactics and behaviors they obey. These 5 types of agents are explained In a nutshell. NegoTo agent, before reaching deadlock applies trade-off tactic, followed by negoEngin tactic. Random agent chooses the next tactic randomly (e.g. negoEngin, trade-off, trade-off and negoEngin). The above-mentioned agents alter negoEngin and trade-off tactics one at a time. TOagent obeys trade-off tactics when utility of new offer is higher than the previous one, otherwise, the aspiration level is decreased and new offer is proposed. Nego agent only follows negoEngin tactics during the negotiation.

We think that applying a decision making model that discovers the highest utility among all tactics can increase the performance of the system. In addition, applying the previous history or knowledge extracted from past experience (e.g. what happened in the past when we used a specific tactic in the same situation?) in the same situation can help agents decision making. This will result in a more accurate choice of tactics or type of agents to negotiate and end in a higher chance of reaching agreement.

C. Attitude of Agents Involved in Negotiation

Internet is populated by many agents coming from different sources with different attitudes and goals. Some researchers mentioned that, although agents can be categorized by their behaviors, they can also be categorized according to their attitudes toward their goals. In a clear word, agent's attitude defines how an agent selects beneficial opponent in a given different situation[17]. Many researchers are working on this area in order to reach to a better efficiency in their models [10, 15, 17, 18] . In Ikpeme proposed model [15] agents are divided according to their social attitudes into three different groups so-called helpful, reciproactive, and selfish. Helpful agents are those who are willing to help. This group benefits in the homogenous groups of helpful agents. Reciproactive agents evaluate the request and will accept opponent's request if the agent meets a balance of cost and saving, otherwise they reject the request. This group has the best performance when situated in an open group of agents. Selfish agents never help the other agents. So, they always benefit when situated among helpful agents but rarely benefit from reciproactive agents. Finally, the results show that in an open environment success rate is many times better for reciproactive agents than selfish agents.

In the work presented by Jaesuk *et al.* [17], agents attitude defines the priority that an agent places on various choices it may have regarding member selection. So, agent's attitude is (1) toward rewards or (2) toward risk. Agent attitude toward reward is the agent's point of view toward finding the best opponent. The attitude considers the opponent's quality of service. Agent attitude toward risk is agent sensitivity to the possible risk of opponent agent (which depends on unreliability and unavailability of agents). The result of the research shows that agents with strong attitude toward risk are more beneficial when there is a higher chance of failing jobs due to tight time or low availability and reliability of opponent. But, agents with the high attitude toward quality situated in a case where time is enough and the penalty value is small are able to earn more benefit. Also, in some cases agent's classification is done according to the agent's behavior as proposed by Ikpeme et al. [15].

The proposed classifications will clear the roadmap for developers in establishing new research in the area of negotiation agent systems. Once developers define the agent classification schema, automatically dimensions and scopes of research area are clarified. Simply, by following previous works conducted in the specified class, the gaps and disadvantages of research area will be highlighted. This will assist researchers to develop and maintain innovative research in the area of negotiation agent models.

IV. Popular Methods of Optimizing Negotiation Systems

Designing an optimized agent negotiation system is one of the important issues addressed recently by many researchers [16, 19-23]. In recent years, a great deal of effort has been devoted toward optimizing negotiation agents[24-27]. Mostly this aims at improving agent i) adaptability, ii) intellectuality, iii) applied strategies, and iv) gathering information.

A. Adaptability

Negotiation agents are situated in open environment where new agents may come and some agents may leave the environment. Agents are characterized by deadlines, volatile preferences, and incomplete information. In such situations, agents must survive by changing their strategies, preferences and even learning opponent's preferences and behaviors. This attempt to change agent's behaviors, preferences, and strategies toward reaching an agreement is called adaptability. This method aims to find the highest satisfactory offer for agents as well as being acceptable for opponent agent. Many researchers [6, 27, 28] applied this methods for optimizing negotiation systems, as it is effective and necessary to assist the autonomy aspect of agents. Zhang *et al.* [27] proposed a new adaptive negotiation strategy which assists negotiation models by enhancing the adaptability of agents. In this scenario, an agent uses adaptive strategy to estimate the strategies and preferences that opponent agents used in the last few offers in negotiation. So, an agent can choose appropriate negotiation factors to adjust its strategy. Even after choosing the strategy, there is a chance to change the parameters of chosen strategy to adapt dynamically to the strategy of another agent. Also, Raymond [6] proposed a negotiation mechanism in which an agent can adapt itself by changing its preferences and behavior using dynamic models. This learning model is applied by using Genetic Algorithm (GA) in process of decision making, so that agent can gradually acquire appropriate negotiation knowledge based on its negotiation history with that opponent. Thus, Raymond negotiation model aims to reach the same property but by using machine learning techniques.

B. Intellectuality

As intellectual capability is one of the substantial characteristics in agent technology, learning appears as one of the most important and effective methods in improving negotiation agents. Recently, many researchers enriched their negotiation agents and models by using different kinds of learning algorithms or machine learning techniques [6, 9, 26, 29-32]. Former attempts in this area were applied by using some sort of machine learning techniques (e.g. GA, fuzzy, re-enforcement learning, simulated annealing and data mining), distribution probabilistic analysis or by applying some pre-

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

programmed strategies. In Intelligent Trading Agents(ITA) model [9] four different types of strategies were introduced to help agent's to choose the next action to be taken. E-Commerce Negotiation (E_CN) [10] also presents complex negotiation tactics combined with a sort of distribution probabilistic analysis to predict the opponent's type which will end to choose the best possible strategy for the next round. Other researchers tried to employ machine learning techniques directly to the agent's decision making engine. For example, Raymond [6] applied a genetic algorithm on the space of possible offer for agent *A* to propose to agent *B*. Fig. 1 represents an example of how agent offer encodes to the chromosomes and draws a tow-point crossover operation according to Raymond Lau[6].

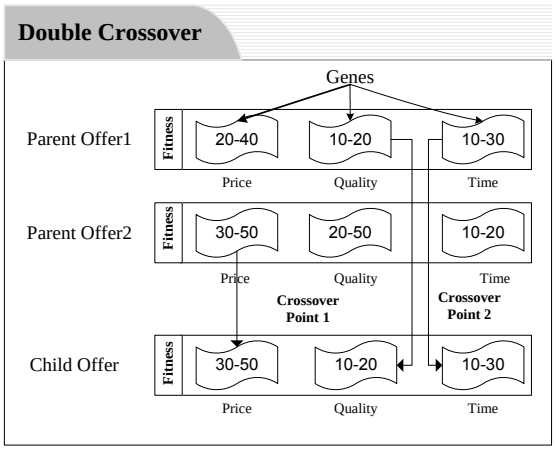


Figure 1. Encoding Candidate Offers[6]

In this scenario agent *A* will generate a random pool of possible offer to propose to agent *B*. Then GA is applied on mating pool which is using three standard genetic operators: cloning, crossover, mutation. The outcome is a list of generated possible offers which are ranked by their utility values. The first offer in the list is the offer to be proposed to agent *B* by agent *A*. Proposed GA-based negotiation agents assist agents by learning their opponent's preferences. This improves the speed of the search in finding the best possible offer for both agents. Although this learning ability increases the overall utility rate to 10.3%, still there are some open issues to increase the performance of the system. One of the open issues is to acquire knowledge extracted from the recorded history or a trusted third party in order to setup the GA initialization accurately (e.g. possible range of value for genes).

Guo *et al*. [32] suggested an algorithm to extract knowledge from a user and then inject it into the solution population. Then, a simulated annealing was applied to render the solution to make sure that the best possible offer will be proposed.

Isabel *et al.* [26] proposed a multi-agent market simulator. In this market, agents have the possibility to negotiate through the pool which is regulated by a market operator (market administrator). Three data mining techniques are proposed to mine the administrator's transaction history, which contains all previous interactions and transactions among agents. The Tow-Step clustering algorithm is used to cluster buyers according to agent's characteristics. After that, a rule-based modeling technique, C5.0 classification algorithm [26] is used to extract the consumption pattern of each cluster population. Finally, Apriory algorithm [26] is used to discover association between buyer details and purchases. A future direction for this research is to employ some sense of artificial intelligence to the seller agent which will result in a higher overall performance of the system.

C. Applied Strategies

The outcome of negotiation depends on several parameters including agent's strategy. Strategies are divided into three groups: (1) strategies which depend on time called time-dependent strategies, (2) strategies which depend on agent's behavior, called behavior-dependent strategies and (3) strategies which depend on how a specific resource is consumed. Recently there have been huge amounts of research on agent's strategies [9, 10, 26, 27, 29].

In ITA model [9], some time-dependent strategies were proposed, namely Desperate, Patient, Optimized patient and strategy Manipulation. Desperate strategy accepts the first acceptable offer that is suitable. In this strategy agent aims to reach a deal as soon as possible. Patient strategy waits until all negotiation threads reach a deal and then choose the best offer. This strategy guarantees the best possible deal but does not consider time constraint which is one of the most important factors in real market places. In Optimized patient strategy, the outcome of one sub-negotiation affects the performance of other sub-negotiations. Manipulation is a combination of above-mentioned strategies. A drawback of such strategies is the lack of adaptability as they are preprogrammed.

E-CN model [10] used time-dependent strategies called Conceder, Linear and Tough. Conceder strategy means that an agent quickly lowers its preference values until it reaches its minimum reserved value. Linear strategy is when agent drops its reserved values but in a gradual manner. Tough strategy deals in a tough manner, meaning that it keeps its values until agent is close to deadline then suddenly drops the values to its reserved values. However, these sorts of strategies are sub-optimal in which using a sense of learning can improve the efficiency and robustness of system.

Isabel et al. [26] proposed two types of behavior-dependent strategies named Composed Goal Directed (CGD) and Adaptive Derivate Following (ADF). CGD is based on two objectives which should follow sequentially. The first objective is to be sure that all needed goods are sold or purchased. The second tries to reduce the pay off of the deal or increase the benefit. ADF is based on the revenue earned in the previous period as a result of changes in the price. If the change of price by the last period produced more revenue than the previous period, then the strategy takes a similar action otherwise it will take a counter action.

Although, time-dependent strategies seem simple, with concern to the time dependent of negotiation agent, they have a significant effect on system's outcome. A combination of alternatives of different types of tactics were proposed in this area [16, 23].

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

D. Gathering Information

The lack of information about environment and opponent agents is one of the major issues on negotiation systems [18, 26, 32-35] since these information assist the agents to choose suitable agents and strategies. Such information can be gathered via agent's recorded negotiation histories or via trusted third party agents and referral systems. The referral mechanism allows agents to find their required resources if there is any agent with the required expertise close to the location of the neighboring agents [35]. Many researchers [18, 26, 32-36] have focused on this issue with the aim to improve their negotiation agent models. Guo [32] in his proposed algorithm referred to information as an important factor which assists suggested model in learning multi-attribute utility function. Gathering information is mentioned as the second step of the Guo algorithm Fig. 2.

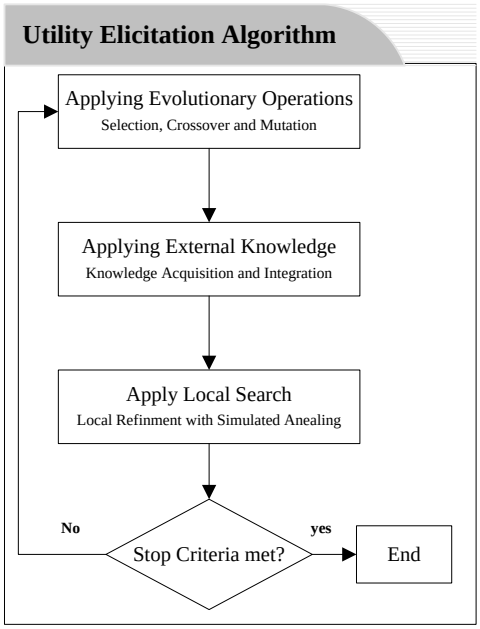


Figure 2. Utility elicitation algorithm[32]

In that second step called "applying external knowledge", knowledge gathered from outside is obtained directly from user input or derived from user observation. The base solution is now modified by injecting external knowledge into the solution population. The knowledge will be accepted based on correctness of the knowledge and a certain probability [32] . Guo et al. [32] showed that taking this knowledge into account while generating solution population, has a significant effect on learning user preferences in multi-attribute negotiation.

Choo et al. [33] also presented a form of optimizing negotiation agents by employing genetic algorithms. This model attempts to learn the agent opponent's preferences according to the history of the counter offers based upon the stochastic approximation. One of the further discussions in the agent models is employing trust on negotiation agents

Trust can help agents to follow their aims by gathering useful information about their opponent's preferences and their attitudes toward their goals. Multi Dimensional Trust (MDT) introduced by Griffiths *et al.* [37] opened a future direction in the area of negotiation agent for finding trustworthy opponent. Griffiths *et al* [37] applied a weighting factor concept which enables agents to combine decision factors according to agents current preferences. Later on, Gujral [38] provided its agents with a model of recognizing the best trusted third party to obtaining information from. In that model, agents should consider the trustworthiness of the potential opponents in order to maximize the agent's rewards, inasmuch as the more reliable the trusty agent is, the higher the chance of reaching agreement.

There are some other important factors in multi agent negotiation systems which, taken into account, can affect the outcome of negotiation's systems. These include the number and type of issues considered in agent services [16, 39], agent attitude [17], one-sided or two-sided commitment [11, 40], bilateral [39] or multilateral [6] negotiation.

The standout research and relative optimization techniques and methods of negotiation agent models are summarized in table 1 below.

TABLE I. POPULAR TECHNIQUES AND METHODS FOR OPTIMIZING AGENT NEGOTIATION SYSTEMS

Optimization Methods	Techniques	Exempular description	Stand out Researches
Gathering information	Referral Systems	Asking information from neighbors	Ebadi et.al. (2008)
	Trust	Based on trust worthy of opponent	Griffiths (2005),Gujral et.al.(2007)
	History	based on recorded experienced	Choo et.al.(2009), Guo et.al.(2003)
Adaptability	adaptive learning techniques	learnin adaptive factor during trading	Raymond (2009),Magda et.al.(2009)
	Flexible models	overcoming pre-programmed tactics	Zang et.al(2007),chung-wei(2008)
	Dynamic models	Dynamic methods and programming	Raymond (2009)
Intellectuality	Marchine learning techniques:	Genetic Algorithm	Choo et.al.(2009), Magda et.al.(2009)
		Bayesian Learning	Duong e.al.(2004),Choo et.al(2009)
		Reinforcement Learning	Sen et.al.(2007)
		Fuzzy Logic	Cheng et.al.(2005)
		Evolutionary Learning	Raymond(2009)
	Data mining techniques	Apriory algo/Classification(C5)Algo	Isabel et.al.(2008)
	Others	Simulated Annealing	Isabel et.al.(2008)
Applied Strategies	Time dependant tactics	Boulware/Conceder/Linear	Iyad et.al.(2002),Doung et.al.(2004)
	imitative tactics	Relative tit-for -tat/Random absolut tit-for-tat/Averaged tit-for-tat	Isabel et.al.(2008)
	Resource dependant tactics	Dynamic deadline/Resource estimation	not common

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

V. APPLICATIONS

After specifying the domain of classification and optimization of negotiation agents, some outstanding works carried out during recent years will be reviewed. Following the new generation of the models we will track the progressing flow of agent generations. In addition, the whole negotiation systems proposed in the models will be analyzed to elaborate on advantages and drawbacks of these systems. These systems are analyzed in terms of desirable negotiation protocol, negotiation strategy, agent characteristics and negotiation setting.

1) Desirable negotiation protocol represents the rules that govern the interaction between negotiators [22]. According to Nguyen and Jennings.[41], desirable properties for a negotiation protocol include pareto efficiency and Guarantee of success.

2) Negotiation strategy is the arrangement of a series of actions that the agents plan to take through the negotiation. Negotiation tactics specify whether it is time dependent, resource dependant or imitative. In some cases, negotiation tactics are a combination of aforementioned tactics proposed by models.

3) Agent characteristics specify agent's knowledge and experience, learning and adaptation capabilities.

4) Negotiation setting deals with factors which are related to problem domain. It includes number of negotiation issues and number of parties involved.

The analysis will deal with the evolution that the primary Kasbah model [7] underwent from 1996 to 2009, leading to GAMA model (table2)

Kasbah model [7] is a simple one-to-one negotiation framework. Its application is on e-commerce where the agent technology will meet the web-based systems and try to overcome the need of online trading by applying some autonomy on trading. Kasbah model is single issue and considers price as the most important issue in negotiation. Therefore, agents are searching to make a deal with appropriate price even before they meet their deadline. Lack of adaptation and intellectuality is obvious draw-back of Kasbah's negotiation agents which is overcome by defining new versions of Intelligent Trading Agents such as ITA and e-CN.

ITA [9] is a one-to-many negotiation framework which is an improved version of negotiation systems in terms of number of issues considered in negotiation and in terms of communicating as it follows bilateral negotiation. In bilateral negotiation, agents have the ability to send offer and counter-offer in both ways. ITA presents new system architecture represented in Fig. 3.

ITA buyer agent includes two components, namely coordinator and sub-buyers. Buyer agent will establish several one-to-one negotiations between sub-buyers of buyer agent and sellers. In this work, agent preferences are represented as a constraint satisfaction problem as described by Vipin [42]. Moreover, this model proposes different sort of strategies like desperate, patient and optimal patient. Optimal patient helps agents to assist agent's autonomous behavior in dynamic environments. In this case, preferences of agents will be marked by weighting factor which represents the degree of importance of every issue.

Weighting factor was firstly introduced by Griffiths [37]. Later on, this concept under the term of Multi-dimensionality [38] was used by many agent researchers [43, 44] to evaluate their standard measurements and finding the more appropriate deal or opponent. However, ITA's strategies help agent to gain higher performance but still the act of agent is bounded by premier choice of strategies before each round of negotiation. This will reduce the adaptability of negotiation agent.

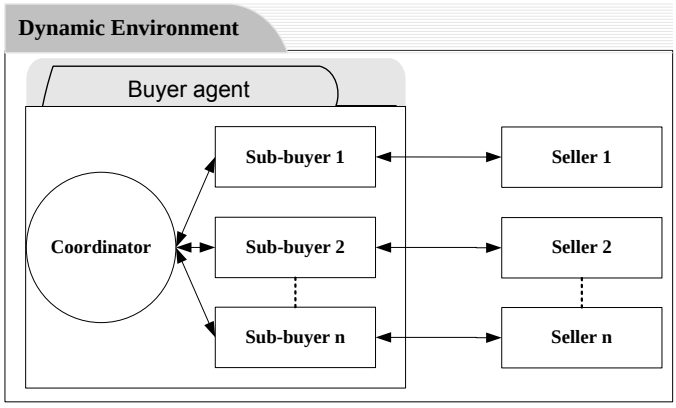


Figure 3. System architecture adapted from ITA system [9]

The new version of ITA called e-CN was proposed by Nguen and Jennings [10]. This method uses several number of agent services' issues considered in negotiation. Also, a factor is introduced as "probability of agent distribution" which represents the probability of allocating types of agent in the environment. Nguyen [10] believes that agents can be divided by their behavior into two groups of agents called Conceder and Non-conceder. Conceder stands for agents who are willing to concede in order to reach a deal while Non-conceder stands for agents who are not willing to sell, otherwise there is a special amount of benefit for them, so they act in a tough manner.

In e-CN model, agents choose their strategies based on a method of predicting the expected utility of chosen strategy considering the current situation. This Expected utility will be evaluated according to 3 important probability factors: probability of agent's distribution, possibility of reaching agreement and average utility value if agent reaches the agreement.

The disadvantage of proposed concession strategies in e-CN and ITA is that in every round of offering new proposal, agents are conceding while there is a possibility to find a mutually acceptable proposal with the same utility level but different values of the issues. Future research should be done

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

to investigate a sort of similarity function that reveals the negotiable issues with higher importance. Finding important negotiable issues assists agents to reach to an agreement using the slightest possible amendments. We believe this will result in higher optimization and also faster deals in terms of time consuming. Aforementioned future developments will result in decreasing negotiation cost.

Although these models support bilateral multiple concurrent negotiation, there is still much more need to assist the agent decision making methods. Helping agent to predict its opponent next action or finding the opponent preferred issues is another future work. As we discussed in section IV this objective could be achieved by enhancing the intellectuality of agents. Referring to table 1 there are many possible techniques to investigate as future studies. An appropriate learning technique could lead this model to find the pareto optimal solution.

Cheng [23] proposed a heuristic, multiple issue, one-to-many model for negotiations in a third-party-driven e-market place. These negotiation agents employ trade-off tactics using fuzzy inference systems to generate new offer in each round. Trade-off tactics are navigated from time-dependent and imitative tactics.

Although, Cheng's proposed model is pareto-efficient and highly adaptive revealing the importance level of issues to the other agents, it violates the privacy of information. These sorts of assumptions are highly inappropriate since they are hardly acceptable in the real world environments such as e-markets.

A negotiation meta strategy for one-to-one bilateral negotiation was proposed by Ros and Sierra[16]. Meta strategy is a combinatorial sequence of concession and trade-off tactics which will try to keep the aspiration level, otherwise there is no possible offer by the current aspiration level. Combining tactics allows agents to outperform better in different situation which fulfills the adaptive capability of agents. In addition, detection of opponent agent preferences helps agents to propose mutually acceptable offers. As a consequence, the system's final success rate increases.

Although this model is placed in a satisfactory level of adaptability, it is recommended that further research be undertaken in order to learn the opponent agent's type. Finding the correct classification for type of agent could increase the chance of reaching agreement.

Choo *et al.* [19] conducted a research on one-to-many bilateral negotiation with multiple issues (quantitative issues). The system architecture of the study was based on ITA's system architecture. They investigated two different machine learning approaches genetic algorithm and Bayesian learning called GA improved-ITA and Bayesian improved-ITA. The result obtained from the final analysis showed that GA-improved-ITA outperforms, Bayesian-improved-ITA in maximizing the joint utility and negotiation payoff at the same time as it increases the negotiation cost.

Raymond [22] performed a similar series of experiments in terms of models characteristics. Raymond negotiation agents are enhanced as they are promising in supporting real-world e-market places environment. This application sustains multi-party, multi-issue, many-to-many negotiation which are based on parallel and distributed decision making model. Moreover, Raymond [22] introduced his novel genetic algorithm in this experiment. The final result of the experiment showed that an evolutionary negotiation agent guarantees pareto optimal solutions underneath dynamic negotiation situation, for example in the incidence of time limitations.

However, as we mentioned in section IV D, embedding a sort of strategies could enhance the success rate of the two aforementioned negotiation systems.

Magda [45] introduced an agent mediated shopping system called Genetic Algorithm driven Multi-Agents (GAMA). GAMA is a multi issue bilateral negotiation system enhanced by learning ability. GAMA studies the effect of participating opponent agent's preferences in decision making of agents. In order to do that, one of the offers from the opponent agent's previous offers (or list of offers) is chosen as one of the parents and the other parent is chosen from agent's own preferable proposals. Then, the mutated offspring is generated. The new generated offspring is a potential satisfactory offer as it is mutually acceptable for both agents. Experimental results demonstrated that GAMA achieved to a higher satisfaction rate while reaching to the higher numbers of the deal in comparison with traditional GA methods.

One of the advantages of this work considering these parent selections is increasing the adaptability of the system, since every change of opponent agent's preferences effects the decision making of agents. In addition, proposing mutually admissible offer causes to reach an agreement in fewer rounds of negotiation, thus reducing the cost of negotiation. In future investigations, it might be possible to experiment this model under qualitative issues as an alternative of quantitative issues. This could result in higher optimization and also faster deals in terms of time consumption.

In order to have a better understanding on the overview of applications discussed, table2 illustrates the general characteristics of the applications discussed.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

TABLE 2. A COMPREHENSIVE ANALYSIS OVER SYSTEM APPLICATION PROGRESS DURING THE LAST THIRTEEN YEARS.

Negotiation system characteristics	Kasbah(1996)	ITA(2001)	e-CN(2004)	Cheng et.al.(2005)	Ros et.al.(2006)	Choo et.al.(2009)	Raymond(2009)	GAMA(2009)
Cardinality of Negotiation Domain	single-issue	multiple-issue	multiple-issue	multiple-issue	multiple-issue	multiple-issue	multiple-issue	multiple-issue
Cardinality of Communication	unknown	bilateral	unknown	bilateral	bilateral	bilateral	mutilateral	bilateral
Number of agents involved in negotiaion	one-to-one	one-to-many	one-to-many	one-to-many	one-to-one	one-to-many	many-to-many	one-to-many
Qualitative negotiation value of issues	–	–	√	√	√	–	√	–
Quantitative negotiation value of issues	√	√	√	√	√	√	√	√
Privacy of model	√	√	√	√	√	√	√	√
Privacy of information	√	√	√	not-private.see Des	√	√	√	√
Negotiation tactics includes:								
TIme/REsource dependant or IMitative	–	Time dependant	mix(TI+IM)	Trade-off(TI+IM)	Mix(TI+IM)	Time dependant	Time dependant	–
Intellectuality	–	–	Bayesian learning	Fuzzy inference	–	Genetic Algorithm, Bayesian learning	Genetic Algorithm	Genetic Algorithm
Pareto efficiency	–	–	–	Preto Optimal	–	–	Preto Optimal	–
Guarantee success	–	√	√	–	–	–	–	–
Type of agents involved	unknown	unknown	conceder non-conceder	NegoTo, Random, aLternate, TOAgent, NegoAgent	unknown	unknown	unknown	unknown
Adaptiveability	non-adaptive	non-adaptive	semi-adaptive	adaptive	adaptive	non-adaptive	adaptive	adaptive

This table shows that embedding strategies in negotiation agent models can increase the final system outcome. This enhancement could be in terms of increasing adaptability or assuring the guarantee of success. As illustrated, most of the models empowered by machine learning techniques are fully adaptive. As discussed in section IV A, adaptability is an important characteristic for negotiation system embedded in open environment. Also, as we can see in e-CN and Cheng *et al.* models [10, 23] an accurate classification on agent's type is driving the negotiation model to a desirable level of adaptability.

In order to reach to the pareto optimal result, agents most be equipped with an effective learning method or a suitable strategy. These cases show that learning ability helps agents to predict opponent's characteristics (e.g. preferences, reservation value, attitude and type) accurately. Therefore, by choosing the best possible strategy (action) we can reach to a pareto optimal result.

Every negotiation mechanism is desirable to meet some important requirements. These include pareto efficiency and guarantee of success. The power of applying suitable strategy is revealed by assuring the guarantee of success only by employing the appropriate strategy.

We believe tables 1 and 2 provide a good resource for future developments. Since table 2 highlights the shortcomings and loopholes of the above-mentioned models (represented by "-" or "non"), developers can easily investigate future studies in order to overcome these gaps by the help of the optimization methods introduced in table1.

As this study shows, artificial intelligence is an important characteristic of multi agent systems which has been used as a popular enhancing technique to optimize the final outcome of many negotiation agent based systems. As Jennings [46] admits" undoubtedly the main contributor to the field of autonomous agents is artificial intelligence".

Artificial intelligence techniques such as GA [6, 19, 32] , fuzzy logic [23, 25], simulated annealing [32] ,neural network [47, 48] and re-enforcement learning [30] have been used broadly to improve agent's intellectuality in recent years. According to our review, GA is the most popular learning technique among others as usually models that employ GA in negotiation systems reach to a higher final utility than other techniques (e.g. [22, 49]) as shown in Fig. 4.

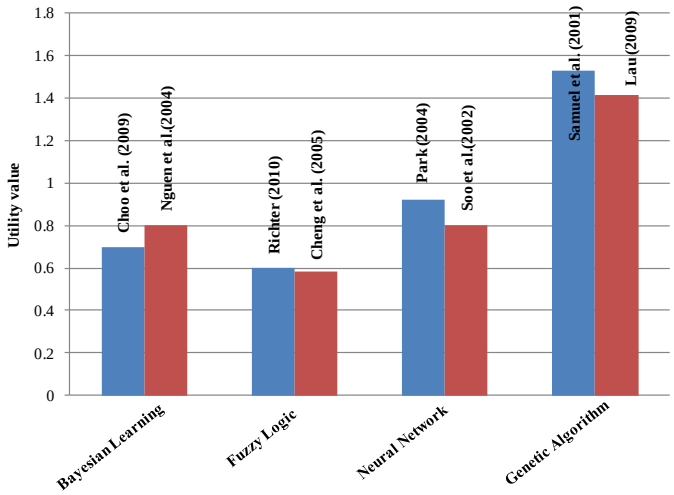


Figure 4. Comparison of final utility value of different machine learning methods

This is due to the pattern of common problems defined in negotiation which match GA technique characteristics, since, in typical negotiation problems, we face a big space of feasible solution and the goal is to find the best solution. Furthermore, specifying the fitness function for population ranking is a straightforward task as generally it equals to utility function for evaluating offers.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

VI. CONCLUSION

In the last few years we have witnessed an incredible growth of multi agent systems on e-commerce. One of the most important driving forces behind MAS research and development is the internet. By increasing the human demand on online trading, negotiation has become one of the most important topics in MAS. In this area, many researchers tried to improve the performance of the negotiation agent's model. This is achieved by using adaptive methods, employing learning abilities, applying strategic reactions and gathering information. To sum up, negotiation agents can act more efficiently when they are empowered with effective methods for gathering information which assists the agents in employing strategic actions to reach their goals. In addition, agents must be adaptive by changing their attitudes and learning their opponents' characteristics and preferences to improve the overall performance of the system.

In order to improve agent negotiation systems, we need to understand the dimensions and range of options in these areas. To set up the foundation, we have developed a classification scheme which is specially aimed at negotiation agent systems on e-commerce. Negotiation agents can be categorized into different groups with respect to (1) number of agents involved in negotiation, (2) type of agents involved in negotiation, and (3) according to negotiation agent attitude involve in negotiation.

This classification system was demonstrated on an assorted range of outstanding negotiation model and the outcome is summarized in table 2. The purpose of this classification was to present a complete and systematic source to objectively compare and contrast different negotiation models. Such a classification method is essential for developers as it supplies a resource of differentiating competing alternatives for the area of negotiation agent's models to exploit.

VII. REFERENCES

[1] M. Wooldridge and N. R. Jennings, "Intelligent Agents: Theory and Practice, " Knowledge Engineering Review, vol. 10, pp. 115–152, October 1995.

[2] C. Castelfranchi, "Guarantees for autonomy in cognitive agent architecture, " Proceedings of the workshop on agent theories, architectures, and languages on Intelligent agents, vol. 890, pp. 56-70, 1995

[3] S. P. K. Michael R. Genesereth "Software Agents, " Communications of the ACM, vol.37, pp. 48–53, 1994

[4] K. P. Sycara, "Multiagent Systems, " AI Magazine vol. 19(2), pp. 79-92, 1998.

[5] M. Woolridge and M. J. Wooldridge, An introduction to multi agent systems,10 ed., John Wiley & Sons, Inc. New York, NY, USA 2001,

[6] R. Y. K. Lau, "Adaptive negotiation agents for e-business, "Proceedings of the 7th international conference on Electronic commerce, pp. 2005

[7] C. Anthony and M. Pattie, "Kasbah: An agent marketplace for buying and selling goods, " In Proceedings of the first international Conference on the Practical Application of Intelligent Agents and Multi-Agent Technology, pp. 1996

[8] V. B. Ryszard Kowalczyk "On Constraint-Based Reasoning in e-Negotiation Agents, " Lecture Notes In Computer Science, vol. 2003, pp.31-46, 2001

[9] I. Rahwan, R. Kowalczyk and H. H. Pham, "Intelligent agents for automated one-to-many e-commerce negotiation, "Proceedings of the twenty-fifth Australasian conference on Computer science , vol.4, pp.197-203, 2002

[10] T. D. Nguyen and N. R. Jennings, "Coordinating Multiple Concurrent Negotiations, "Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems , vol.3, pp.1062-1069, 2004

[11] J. Dang and M. N. Huhns, "Concurrent Multiple-Issue Negotiation for Internet-Based Services, " IEEE Internet Computing, vol. 10, pp. 42-49, 2006

[12] D. Jiangbo and H. M. N., "Coalition deal negotiation for services, " Rational, Robust, and Secure Negotiation Mechanisms in Multi-Agent Systems, 2005, pp. 67-81, 2005

[13] J. Dang, J. Huang and M. N. Huhns, "Workflow coordination for service-oriented multiagent systems, "Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems, pp. 2007

[14] W. Danny, S. Kurt, H. Tom and G. Olivier, "Towards Adaptive Role Selection for Behavior-Based Agents, "Adaptive Agents and Multi-Agent Systems III, pp. 295-312, 2005

[15] I. Erete, E. Ferguson and S. Sen, "Learning task-specific trust decisions, "Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems , vol.3 pp. 2008

[16] R. Ros and C. Sierra, "A negotiation meta strategy combining trade-off and concession moves, " Autonomous Agents and Multi-Agent Systems, vol. 12, pp. 163-181, 2006

[17] J. Ahn, D. DeAngelis and S. Barber, "Attitude Driven Team Formation using Multi-Dimensional Trust, "Proceedings of the 2007 IEEE/WIC/ACM International Conference on Intelligent Agent Technology, pp. 229-235,2007

[18] E. Toktam, P. Maryam and P. Martin, "Partner Selection Mechanisms for Agent Cooperation, " Web Intelligence and Intelligent Agent Technology, 2008. WI-IAT '08. IEEE/WIC/ACM International Conference on, vol. 3, pp. 554-557, 2008

[19] M. N. S. a. M. H. S. S.C. Ng, "Machine learning approach in optimizing negotiation agents for e-commerce, " Information Technology Journal, vol. 8, pp. 801-810, 2009

[20] L. Alessio, W. Michael and J. R. Nicholas, "A Classification Scheme for Negotiation in Electronic Commerce, " Group Decision and Negotiation, vol. 12, pp. 31-56, 2003

[21] P. Faratin, Carles Sierra and N. R. Jennings, "Negotiation Decision Functions for Automated Agents, " Elsevier Sience, vol.24 pp.159-189, 1997

[22] R. Y. K. Lau, "An Evolutionary Approach for Intelligent Negotiation Agents in e-Marketplaces, " Intel. Agents in the Evol. of Web & Appl, vol. 167, pp. 279–301, 2009

[23] C.-B. Cheng, C.-C. h. Chan and K.-C. Lin, "Intelligent agents for e-marketplace: Negotiation with issue trade-offs by fuzzy inference systems, " Decis. Support Syst., vol. 42, pp. 626-638, 2005

[24] M. A. Lopez-Carmona, I. Marsa-Maestre, J. R. Velasco and E. d. l. Hoz, "Using Clustering Techniques to Improve Fuzzy Constraint Based Automated Purchase Negotiations, " Springer, Advances in Agent-Based Complex Automated Negotiations, pp. 89–117, 2009

[25] J. Richter, "Multistage Fuzzy Decision Making in Bilateral Agent Negotiation, " 3sd PHD symposium, pp. 71-73, 2010

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

[26] I. Pra, M. Jo, o. Viamonte, Z. Vale and C. Ramos, "Agent-based simulation of electronic marketplaces with decision support, "Proceedings of the 2008 ACM symposium on Applied computing, pp. 2008

[27] Z. Junyan, T. Jiang and D. Gang, "Agent-based multi-factors adaptive negotiation in E-Commerce, " Grey Systems and Intelligent Services, 2007. GSIS 2007. IEEE International Conference on, pp. 1528-1532, 2007

[28] C.-W. Hang, Y. Wang and M. P. Singh, "An adaptive probabilistic trust model and its evaluation, "Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems, vol.3, pp.1485-1488, 2008

[29] L.-k. Soh and J. Luo, "Combining individual and cooperative learning for multi-agent negotiations, "Proceedings of the second international joint conference on Autonomous agents and multiagent systems, pp.1122-1123 ,2003

[30] S. Sen, A. Gursel and S. Airiau, "Learning to identify beneficial partners, " *AAMAS07* '07 Honolulu, HI USA, 2007

[31] S. Saha, A. Biswas and S. Sen, "Modeling opponet decision in repeated on-shot negotiations, " AAMAS'05, ACM, vol. july,pp. 25-29.2005

[32] G. Yutao, M. Jörg and W. Christof, "Learning User Preferences for Multi-attribute Negotiation: An Evolutionary Approach, "Multi-Agent Systems and Applications III, pp. 1067-1067, 2003

[33] n. s. choo, machine learning approach for optimizing negotiation agents, UniversityPutraMalaysia, pp.167, 2007,

[34] T. D. Nguyen and N. R. Jennings., " A heuristic model for concurrent bi-lateral negotiations in incomplete information settings, " Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence 2003, vol. pp. 1467–1469, 2003

[35] T. Ebadi, M. Purvis and M. Purvis, "Finding Interaction Partners Using Attitude-Based Decision Strategies, "Proceedings of the 2008 Ninth International Conference on Parallel and Distributed Computing, Applications and Technologies, vol. pp. 2008

[36] S. Kraus, "Beliefs, time and incomplete information in multiple encounter negotiations among autonomous agents, " Annals of Mathematics and Artificial Intelligence vol. 20, pp.1-4, 1997

[37] N. Griffiths, "Task delegation using experience-based multi-dimensional trust, "Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems,vol.11, pp. 489-496, 2005

[38] N. Gujral, D. DeAngelis, K. K. Fullam and K. S. Barber, "Modeling Multi-Dimensional Trust, " AAAI07_Materials, vol. 6, pp.35-41, 2007

[39] S. S.Fatima, M. Wooldridge and N. R. Jennings, "Approximate and online Multi-issue negotiation, " AAMAS'07, pp. 947-954, 2007

[40] T. D. Nguyena and N. R. Jennings, "Managing commitments in multiple concurrent negotiations, " Electronic Commerce Research and Applications, vol. 4, pp. 362-376, 2006

[41] P. F. NR Jennings, AR>Lomuscio , S Parsons, M. Wooldridg, C. Sierra, "Automated Negotiation: prospects, method and challenges, " Group Decision and Negotiation, vol.10 pp.199-215 2001

[42] V. Kumar, "Algorithms for constraint-satisfaction problems: a survey, " AI Mag., vol. 13, pp. 32-44, 1992

[43] D. J. Kim, Y. I. Song, S. B. Braynov and H. R. Rao, "A multidimensional trust formation model in B-to-C e-commerce: a conceptual framework and content analyses of academia/practitioner perspectives, " Decis. Support Syst., vol. 40, pp. 143-165, 2005

[44] J. Ahn, X. Sui, D. DeAngelis and K. S. Barber, "Identifying beneficial teammates using multi-dimensional trust, "Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems, vol. 3pp. 2008

[45] M. B. Fayek, I. A. Talkhan and K. S. EL-Masry, "GAMA(Genetic Algorithm driven Multi-Agents) for E-commerce Integrative Negotiation, vol. pp. 1845-1846, 2009

[46] N. R.Jennings, K. Sycara and M. Wooldridge, "A Roadmap of Agent Research and Development, " Autonomous Agents and Multi-Agent Systems, vol. 1, pp. 7-38, 1998

[47] V.-W. Soo and C.-A. Hung, "On-Line incremental learning in bilateral multi-Issue negotiation, " AAMAS'02, vol. pp. July 15.2002

[48] S. Park and S.-B. Yang, "An Efficient Automated Negotiation System Using Multi-attributes in the Online Environment, " lecture Notes In Computer Science, vol. 3140, pp. 544–557, 2004

[49] P. M. C. Samuel, L. Jiming and S.-P. Chan, "Evolutionary Negotiation in Agent-Mediated Commerce, " lecture Notes In Computer Science, vol. 2252, pp. 224–234, 2001

AUTHORS PROFILE

Sahar Ebadi is currently doing her Master degree in Faculty of Computer Science and Information Technology, UPM. Sahar has received her B.Sc in software computer engineering field in 2006 from Iran Azad University. Her research interest includes Artificial Intelligence and Autonomous Agents and Data mining.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Customized Digital Road Map Building using Floating Car GPS Data

G. Rajendran
Assistant Professor of Computer Science,
Thiruvalluvar Government Arts College,
Rasipuram-637401, Tamilnadu, India
guru.rajendran@yahoo.com

Dr. M. Arthanari
Director,
Bharathidasan School of Computer Applications,
Ellispettai-638116, Tamilnadu, India
arthanarimsvc@gmail.com

M. Sivakumar
Doctoral Research Scholar,
Anna University, Coimbatore, Tamilnadu, India
sivala@gmail.com

***Abstract*—The vehicle tracking, navigation and road guidance applications are becoming more popular but the presently available digital maps are not suitable for many such applications. Among the drawbacks are the insufficient accuracy of road geometry and the delayed time in loading the unwanted data. Most of the commercial applications in vehicle tracking require digital maps which have only roads and places of interest whereas the currently available maps show all available roads and places. A simplified map building process to construct customized high-precision digital maps from floating car data obtained from Global Positioning System (GPS) receivers is presented in this paper. The data collected from the GPS receiver fixed in a moving car are used to construct the customized digital road maps. The approach consists of six successive steps: Collecting floating car data (FCD) for desired road segments in a log file; refining the log file; constructing the road segments using the data present in refined log file; merging the segments which has negligible slope; refining the road intersections; and labeling the points of interest. The quality of outcome of the map making process is demonstrated by experimental results and the results indicate that customized road maps of required routes with good accuracy can be built with the help of the proposed map making process.**

Keywords- digital map; global positioning system; floating car data; road network

I. INTRODUCTION

Map is a total or partial depiction of the structure of the earth (or sky) on a plane, such that each point on the map corresponds to an actual point on the earth (or in the sky). Digital maps are used to produce this depiction in electronic form. The digital maps are usually represented as graphs, where the nodes represent intersections and the edges are unbroken road segments that connect the intersections [1]. Each segment has a unique identifier and additional associated attributes, such as the number of shape points that approximate its geometry roughly, the road type (e.g., national highway, state highway, city road, street etc.), name, speed information, and the like. Digital map building is a process that utilizes the information supplied by external equipments in terms of physical features of an environment [2, 3, 4]. This information is taken from different positions along the path followed by a moving object. Once the map is obtained, it can be used to improve the quality of the paths, to locate a target, to help object recognition, to define expectations in the trajectory or to replay the travelled path of a moving object.

Digital maps of required roads with good accuracy are needed in a number of commercial applications. But the presently available digital maps are fully populated with dense roads and other information, most of which are irrelevant to the requirement. Because of the presence of these unwanted data, the loading time of the map in computer memory is also more which results in slower execution of the application. Hence the need for building customized digital road maps is essential and such a map building process will also eliminate the expenses involved in buying digital maps. This paper presents a customized digital road map building process which can be used to construct high quality maps of desired roads and locations. The source code for the intermediate processing steps is written in Matlab 7.6.

Map building process has already been discussed by some of the researchers, but with limitations like complexity in map building and insufficient accuracy. These limitations have been addressed in this work. The remainder of this paper is organised as follows. Section 2 of this paper describes road network model and the collection of floating car GPS data. Related work in this area is discussed in Section 3. In Section 4, the map making process is discussed. The experimental results are dealt in Section 5. The work is concluded and the possible improvements are discussed in Section 6.

II. THE ROAD NETWORK MODEL AND THE FLOATING CAR DATA

A. *The Road Network Model*

The road network data are the basis for vehicle tracking and related applications. The road network model is represented with two-dimensional line segments.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

The road network model [5] is represented by the equation

R_n = (N, S),
N={n | n=(x, y), x, y Є Coordinates},
S= {s | s = <m, n>, m, n Є N},

wherein R_n represents the road network, N represents the node set that indicates the coordinate point set of the road in the road network which is a pair of longitude and latitude (x,y), S represents the road segment set of the road network which is composed by the sequence <m, n>. S represents one directional road that has the beginning node m=begin(s) and the termination node n=end(s). Fig. 1 shows a road network with nodes, segments and an intersection point.

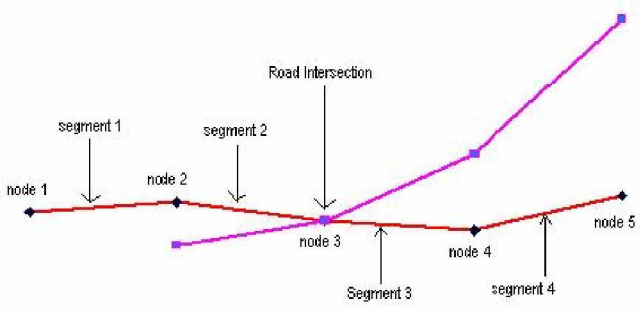


Figure 1. Basic elements of the road network model.

If there is a road segment sequence <s_i, s_j,…, s_k> in the network R_n=(N, S), the termination node of every road segment is the beginning node of the next road segment, this sequence is called one directional road of the road network.

B. Floating Car Data

Nowadays, the main research focus in the community of Intelligent Transport Systems (ITS) is how to acquire real-time and dynamic transportation information. This information can be applied in the transportation area like vehicle tracking, navigation, road guidance and so on. GPS is one of the system which is used to provide real time information on moving objects. It is a Satellite Navigation System which is funded and controlled by the U. S. Department of Defense [6, 7]. The system consists of three segments viz., satellites that transmit the position information, the ground stations that are used to control the satellites and update the information, and finally there is the GPS receiver that computes its location anywhere in the world based on information it gets from the satellites [8].

Figure 2. A moving vehicle (Floating car) fixed with a GPS receiver

The Floating Car (Probe Car) technique is one of the key technologies adopted by the ITS to get the traffic information in recent years [9]. Its basic principle is to periodically record the location, direction, date, time and speed information of the traveling vehicle from a moving vehicle with the data of the Global Position System (GPS) as shown in Fig 2. The information can be processed by the related computing model and algorithm so that the floating car data can be associated with the city road in real time [10]. This data can also be used as a source of data for creating research and commercial applications on vehicle tracking and road guidance systems.

III. RELATED WORK

The history of map making process starts with the Egyptians who for the first time constructed a map for revenue collection three thousand years ago. The digital map building is a new concept developed after the revolution in Information Technology. Though some work has been done in this area, a number of map building techniques are being proposed to suit the emerging requirements.

Y.L. Ip et al., have presented a technique for on-line segment-based map building in an unknown indoor environment from sonar sensor observations [4]. In their approach, the environment data is first obtained from a rotating sonar sensor, analyzed and fed to the Enhanced Adaptive Fuzzy Clustering (EAFC) module to extract the line segments within the workplace of robot. The basic motive is to use full data set to obtain an initial approximation cluster centers via Fuzzy c-mean. This initial approximation helps reducing the number of iterations required for Adaptive Fuzzy Clustering (AFC). This approach is somewhat similar to Fast Fuzzy Clustering (FFC) [11] which is a strategy to speed up the Fuzzy c-mean (FCM). In order to facilitate the map building, the workplace of the robot is divided into squared areas as cells in order to extract the line segments. This mechanism reduces the computation time when extracting the line segments within the world frame. EAFC uses the Noise Clustering (NC) technique proposed in [12] to extract the line segments. EAFC also uses adaptive fuzzy clustering algorithm [13] and fast fuzzy clustering [11]. These algorithms are combined into a single algorithm with enhanced characteristics such as improvement in the computational burden and reduction of the effect of noisy data in fuzzy clustering algorithm. Besides, the authors have proposed a compatible line segment merging technique to combine the similar line segments to a single long line segment as a mechanism to reduce the number of segments in the world model and further improve the quality of the map. This technique is applicable for constructing maps of indoor environment related to robotic applications.

Stefan Schroedl et al., have contributed to map-building by introducing a system that generates digital road maps that are significantly precise and contain descriptions of lane structure, including the number of lanes and their locations, along with detailed intersection structure[1]. The authors combine a large quantity of possibly noisy data from GPS for a fleet of vehicles, as opposed to a small number of highly accurate points obtained from surveying methods. It is assumed that the input probe data is obtained from vehicles that go about their usual

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

business unrelated to the task of map construction, possibly generated for other applications based on positioning systems. The work of authors include the development of a spatial clustering algorithm for inferring the connectivity structure of the map from scratch, the development of a lane clustering algorithms that can handle lane splits and merges and forming an approach to inferring detailed intersection models. This system requires data from hundreds of vehicles already connected to tracking systems for constructing a single segment of the road.

Thus a few number of digital map building techniques are available but they suffer from high complexity of map building process, requirement of more data and dependence on technical skills of the person who is working with map building. The map building process proposed in this work addresses these problems. The process discussed here is a very simple one and even a novice user without technical skills can easily create route maps according to his requirements.

IV. MAP BUILDING PROCESS

A. *Collecting Floating Car GPS data in a log file*

GPS receiver communication is defined with National Marine Electronics Association (NMEA) specification. The NMEA has developed a specification that defines the interface between various pieces of marine electronic equipments. The NMEA standard permits marine electronics to send information to computers and to other marine equipments [14] in predefined formats. Most computer programs that provide real time position information recognize data that are in NMEA format which includes the complete latitude, longitude, velocity and time computed by the GPS receiver. In NMEA specification system, the collected GPS data is converted into a line of text, called a sentence, which is totally self contained and independent from other sentences. The commas act as terminators for the sentences and the programs that read the data should only use the commas to determine the end of a data item.

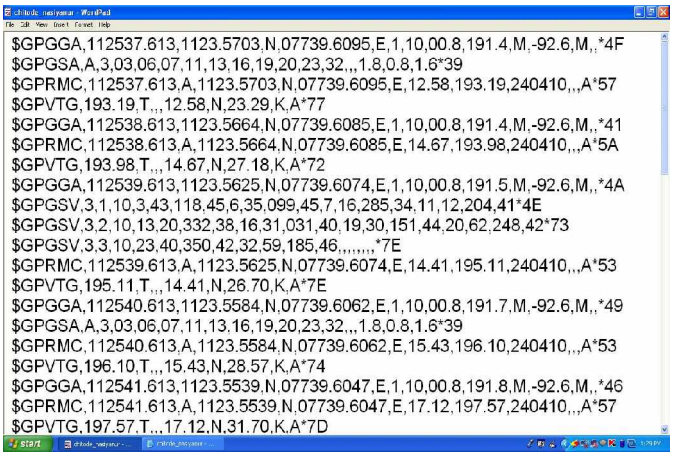


Figure 3. Log file of floating car GPS data with $GPGGA, $GPGSA, $GPRMC, $GPVTG and $GPGSV sentences

In order to collect the floating car GPS data, Wonde-X series GPS receiver (ZX4125) module was used. This GPS receiver module was fixed in a moving car and the generated NMEA sentences are stored in a log file in a laptop kept in the moving car. This GPS receiver generates $GPGGA, $GPGSA, $GPRMC, $GPVTG and $GPGSV sentences at regular time interval of one second. A list of NMEA sentences produced by the GPS receiver and stored in a log file when travelled in a road is given in Fig. 3.

B. *Refining the log file to get $GPRMC sentences*

The log file contains a number of different types of sentences but the recommended minimum sentence C, $GPRMC, provides the essential GPS PVT (Position, Velocity and Time) data. This data is used to locate moving objects in terms of latitude and longitude. The $GPRMC data format is given in Table I. The moving object, if attached with a GPS receiver, can be located with the help of this NMEA sentence.

TABLE I. $GPRMC DATA FORMAT

Data Item	Format	Description
Message ID	$GPRMC	RMC protocol header.
UTC Time (Coordinated Universal Time)	hhmmss.sss	Fix time to 1ms accuracy.
Status	Char	A Data Valid. V Data invalid.
Latitude	Float	Degrees * 100 + minutes.
N/S Indicator	Char	N=north or S=south.
Longitude	Float	Degrees * 100 + minutes.
E/W Indicator	Char	E=East or W=West.
Speed over Ground	Float	Speed Over Ground in knots
Course over Ground	Float	Course Over Ground in Degrees
Date	ddmmyy	Current Date
Magnetic Variation	Blank	Not Used
E/W Indicator	Blank	Not Used
Mode	Char	A Autonomous
Checksum	*xx	2 Digits
Message Terminator	<CR><LF>	ASCII 13, ASCII 10

An example of $GPRMC NMEA sentence is given below:

$GPRMC,120642.206,A,1118.4253,N,07742.4325,E,31.6, 317.52,140510,,,A*62

Where

$GPRMC : Recommended Minimum sentence C
120642.206 : Fix taken at 12:06:42.206 UTC
A : Status A=active or V=Void.
1118.4253,N : Latitude 11 deg 18.4253' N
07742.4325,E : Longitude 77 deg 42.4325' E
31.6 : Speed over the ground in knots
317.52 : Course over the ground

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

140510 : Date – 14th of May 2010
A : Autonomous mode
*62 : The checksum data, always begins with *

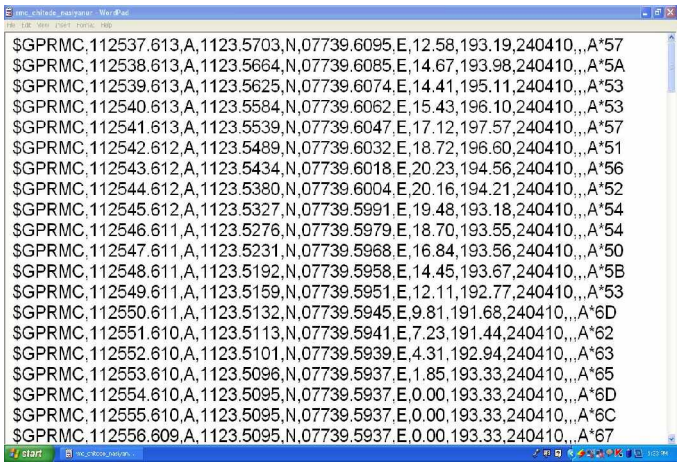


Figure 4. Refined Log file of $GPRMC sentences

Hence the next step in map making process is to refine the log file by removing other sentences in such a way that it contains the $GPRMC sentences only as shown in Fig 4. This refined log file now contains the path of the probe car in terms of latitude and longitude at an interval of one second per sentence.

C. *Segment Extraction*

The road segment extraction is done by considering the locations of the probe car at a fixed time interval. Since in most cases there is no significant distance between the probe car and the road centre line, it is assumed that the path of the vehicle is along the road centre line. The idea is to pick out 'k' locations in road centre line continuously by interleaving 'n' $GPRMC sentences. When this process is iterated, the i^{th} location of the moving vehicle x_i,y_i (longitude x_i and latitude y_i) for all the values of 'i' ranging from 1 to 'k' along the road centre line at 'm' seconds interval is obtained.

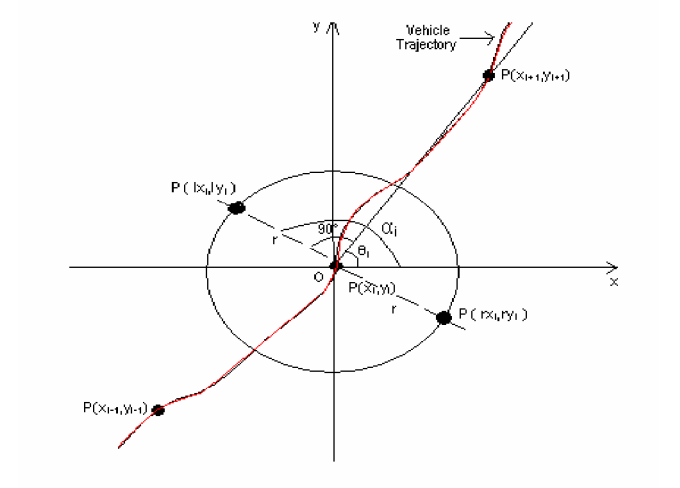


Figure 5. Obtaining Points for Left and Right Road lines using a Point of Road Centre line

After obtaining the points at road centre lines after interleaving m sentences in the middle, the slope between two adjacent points say x_{i+1},y_{i+1} and x_i,y_i is calculated. Now an imaginary line perpendicular to the slope is drawn through x_i,y_i and they are made to intersect an imaginary circle of radius 'r' on both sides as shown in Fig. 5. The points of intersection on either side (lx_i, ly_i and rx_i, ry_i) are recorded and they act as the nodes for the left line of the road and right line of the road, forming road segments. This process is repeated for all values of 'i' ranging from 1 to k-1. It is to be noted that the radius of the imaginary circle determines the width of the road segment. Roads with different widths can be drawn by altering the radius 'r' of the imaginary circle.

The following algorithm is used to extract segments on either side of the road centre line. This algorithm extracts segments by computing road vectors on either side of the road and adds lines in vectors. The output of this algorithm for a sample data is given in Fig 5.

Input : Refined log file; sentence count m; radius of imaginary circle r; number of sentences to be interleaved n.

Output : Line segments on the left and right side of the road forming a road segment; line vector lx, ly for road left line and rx, ry for road right line.

Step 1 : Initialize data items: m=0; i=0; n=40, r=0.5.
Step 2: Repeat step 3 to step 4 till the end of log file.
Step 3: Read the next sentence from the log file
Step 4: If m > =n
  Read the sentence from the log file and
  store longitude into x_i and latitude into y_i.
  i=i+1
  m = 1
 Else
  m = m +1
Step 5 : k = i-1
Step 6 : Repeat step 7 for i =1,2,….,k-1
Step 7 :
  $dx=x_{i+1} - x_i$
  $dy=y_{i+1} - y_i$
  $slope\_radians = \tan^{-1}(dy/dx)$
  $\theta_i = slope\_radians / (\pi/180)$
  $\alpha = \theta_i + 90$
  $left\_plot\_radians = (\pi /180) * \alpha$
  $lx_i = x_i + r*\cos(left\_plot\_radians)$
  $ly_i = y_i + r*\sin(left\_plot\_radians)$
  $Plot(lx_i, ly_i)$
  $\beta = \alpha + 180$
  $right\_plot\_radians = (\pi/180) * \beta$
  $rx_i = x_i + r*\cos(right\_plot\_radians)$
  $ry_i = y_i + r*\sin(right\_plot\_radians)$
  $Plot(rx_i, ry_i)$
Step 8 : line(lx,ly) - Add the line in vectors lx and ly to the current axes.
Step 9 : line(rx,ry) - Add the line in vectors rx and ry to the current axes.
Step 10 : Stop.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

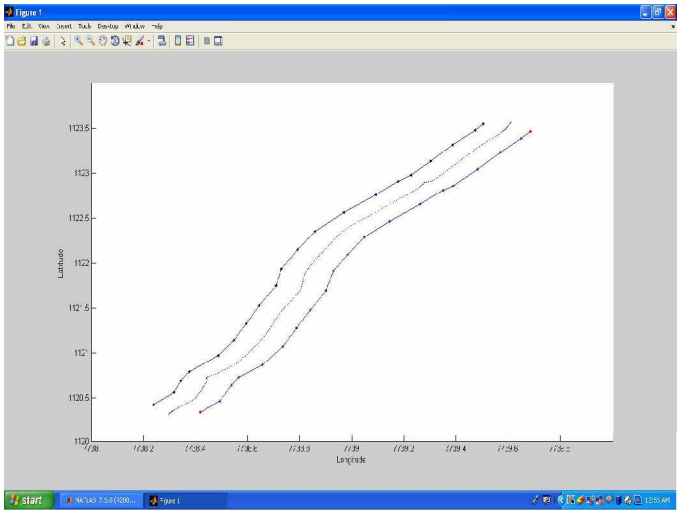


Figure 6. Road segments extracted on the left and right sides of road centre line (vehicle trajectory)

D. *Segment Merging*

The road segment merging technique is used to merge the similar basic road segments together to form a single road segment. It is observed from the outcome of the segment extraction algorithm that a number of adjacent road segments are similar in direction and they can be merged together to form a single segment. The primary advantage of segment merging is that it reduces the number of nodes in the road network and hence simplifies the map. It also results in faster generation of the map because of the less number of segments in the map.

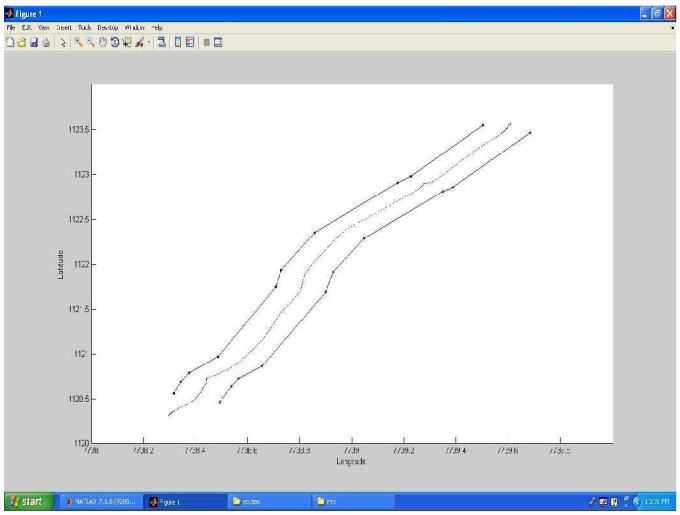


Figure 7. Road segments extracted on the left and right sides of floating car GPS data after segment merging

The criteria used to merge the road segments is the slope (angle) θ_i of the line connecting x_i,y_i and x_{i+1},y_{i+1} with respect to x axis. A threshold limit is set for the negligible slope (angle) λ and if this slope(angle) is within the threshold, say 5° for two adjacent segments, then both the segments are merged together, otherwise the segments are left as they were. This process is iterated for subsequent segments until all the segments are processed.

The algorithm used for segment merging is given below.

Input : Negligible slope (angle) λ, road segments (vectors for left line lx,ly and right line rx,ry).

Step 1 : Compute the slope(angle) of the line connecting x_i,y_i and x_{i+1},y_{i+1} of two adjacent segments (θ_i and θ_{i+1}) with respect to 'x' axis.

Step 2 : Repeat step 3 till all the segments are processed (i=1,2,…,k).

Step 3 : if abs(θ_i-θ_{i+1}) < λ, merge the two segments i.e., remove intermediate node; otherwise do not merge.

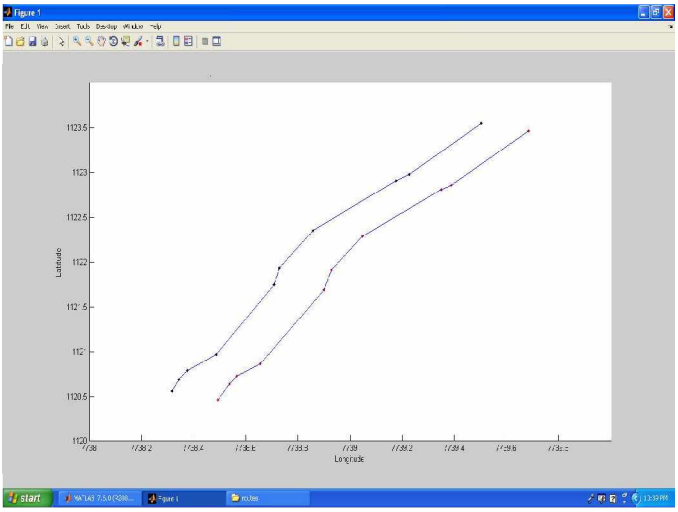


Figure 8. Road segments with left and right road lines after removal of vehicle trajectory

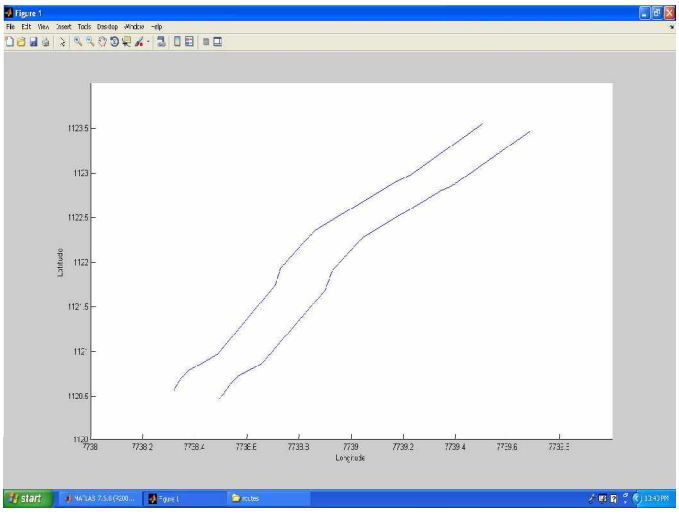


Figure 9. Road Map after removing the plots for nodes

Thus the segment merging algorithm is based on negligible slope and it merges the similar segments together which results in faster production of digital maps by eliminating unwanted nodes in the road network. The extracted segments are shown

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

in Fig. 6 and the merged segments are shown in Fig 7. It is observed that different threshold values for negligible slope (angle) λ can be used to get required accuracy or width of the route map. The merged segments after removing vehicle trajectory are shown in Fig. 8. The final map of the road after removing the plots for nodes is shown in Fig. 9.

E. Refining Intersection of Road Segments

The obtained road network after segment merging is still unrefined at intersections of road segments. The refinement of road intersection is done in the following steps.

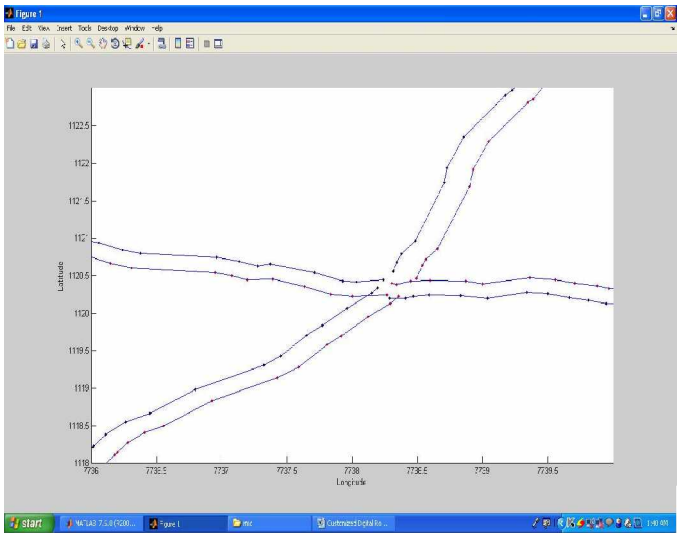


Figure 10. Digital Map with unrefined intersection of road segments

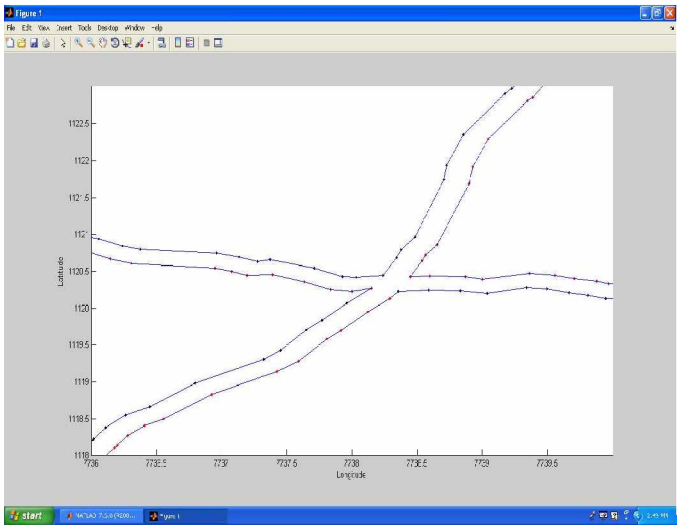


Figure 11. Digital Map with refined intersection of road segments

- Adding a segment: Many a times, there may be a necessity to add one segment either at the end or at the beginning of a road so that the road is connected to another road. In that case, a new segment is added. Sometimes, instead of adding an entire segment, only the left or right line of the road segment is extended.
- Removal of a segment: Sometimes, a segment at the end of one road and a segment at the beginning of the next road may intersect and cross one another. In that case, one of the segments is deleted.
- Extending a segment: When two or more segments do not join at the intersection, the segments are extended based on the previous slope till they form refined intersection. It is to be noted that extending a segment is different from adding a segment.

A map with unrefined road segments intersection is shown in Fig. 10. The road intersection is refined and Fig. 11 shows map with refined intersection of road segments. This process is repeated for entire set of road intersections till a fine-tuned map is obtained.

F. Labeling Points of Interest

A map contains points of interest which means places which are important and noted down on the maps. Placing text on a map is a particularly difficult challenge in digital maps because the axis of the digital maps can be changed dynamically. This is the final step in this map making process and points of interests are noted down from the probe car data and the text is placed so that it is readable and easily located. Care has also been taken that the text does not interfere with the map's data or design. Different font types, styles, sizes, and colors are used to establish clear association between text and map features like telephone post, petrol bunks and toll gates. Fig. 12 displays a legible point of interest.

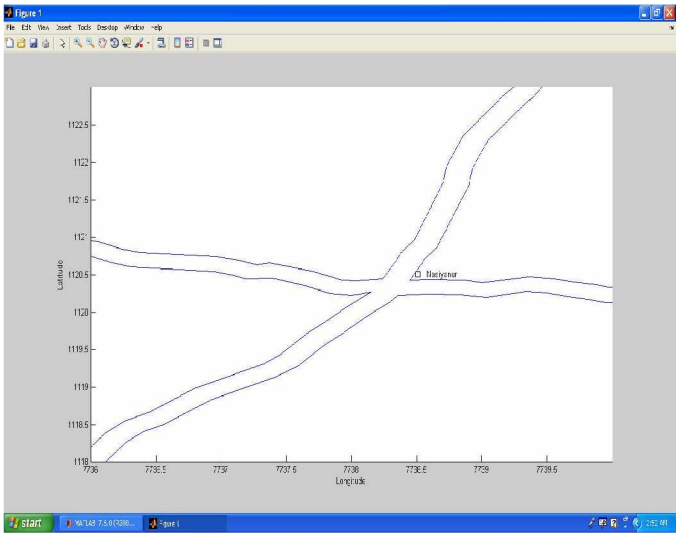


Figure 12. Digital Map with Labeled Points of Interest

V. EXPERIMENTAL RESULTS

The outcome of the map making process is a list of segments drawn on the map with refined intersections of roads. For the purpose of demonstration, the floating car data was collected in different roads. Segments for the roads are extracted using the segment extraction algorithm. Thereafter segment merging was done based on the slope of the adjacent road segments. Finally, the road intersection points are refined and points of interest are labeled to get the final map. The

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

following experiments demonstrate the simplicity and accuracy of the map making.

A. *Simplicity*

A comparison of the map generated by the proposed map making process as shown in Fig. 13 versus a digital map available in the web as shown in Fig. 14 indicates that the proposed process is simple and the map includes only the desired routes and points of interest. The generated map can be easily interpreted because of its simplicity.

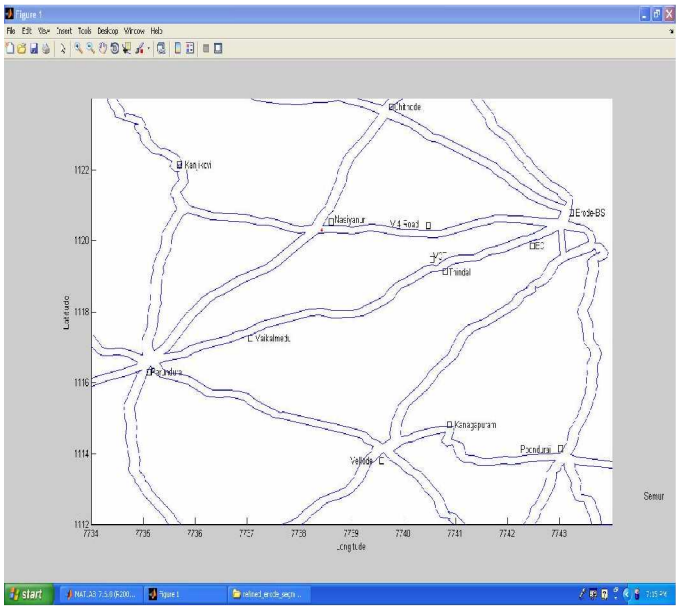


Figure 13. Map generated by the digital map building process discussed in this paper

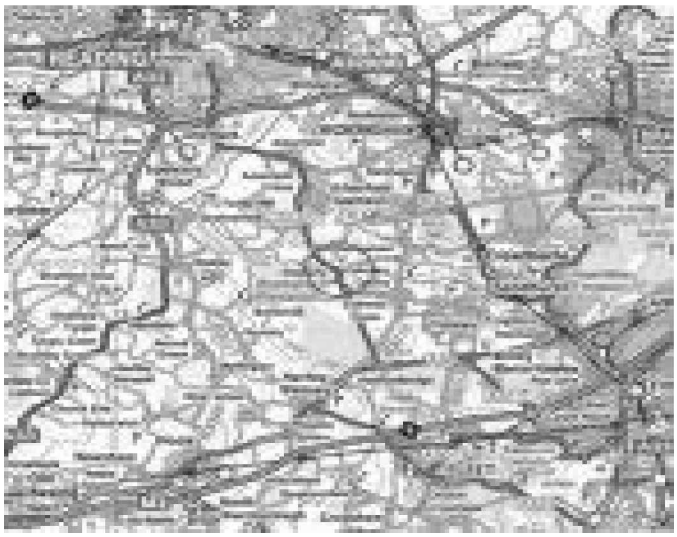


Figure 14. A previously available digital Map with dense routes and places

B. *Accuracy*

One of the input parameter of segment extraction algorithm is the radius of the imaginary circle 'r'. By adjusting the value of 'r' the road segment may be made thicker or thinner to get more accurate segments depending on the necessity. Two different segment extractions for the same floating car data with different values of 'r' according to varying requirement is shown in Fig. 15, 16. It is to be noted that the nodes are shown only to differentiate segments and they are removed in the final map. A more accurate and thinner road is shown in Fig. 16 compared to the one shown in Fig. 15.

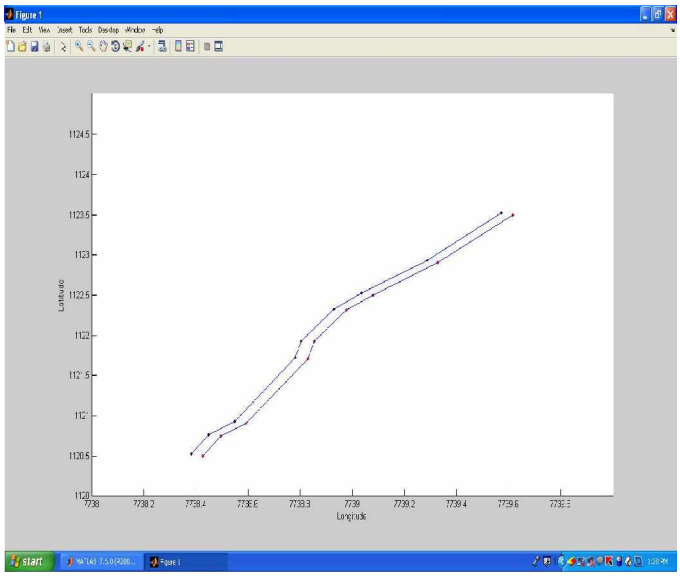


Figure 15. Map with thick roads with less accuracy

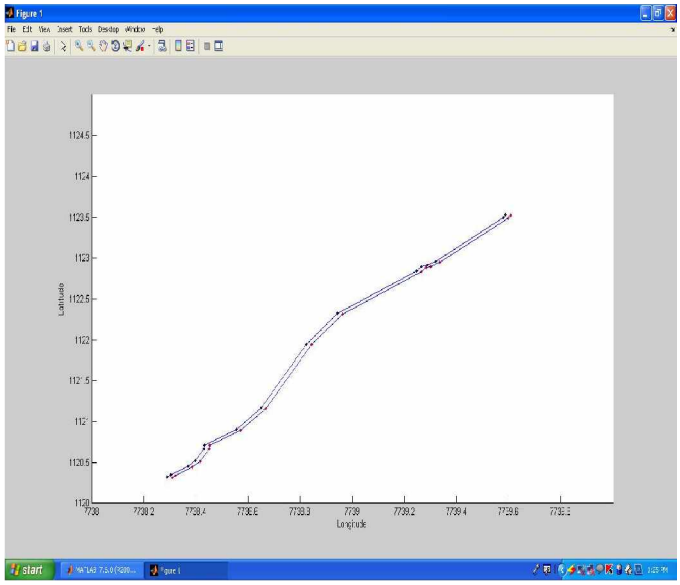


Figure 16. Map with thin roads for the same data with more accuracy

Accuracy can also be obtained by making the sampling sentences interval as minimum. Table II shows a Comparison of segments extracted and segments generated after merging at different intervals of sampling sentences. Accuracy increases with more number of extracted segments and merged segments. The graph shown in Fig. 17 gives a comparison between the

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

number of extracted and merged segments for a floating car data at different sampling intervals.

TABLE II. COMPARISON OF SEGMENTS EXTRACTED AND SEGMENTS GENERATED AFTER MERGING AT DIFFERENT INTERVALS OF SAMPLING SENTENCES.

Interval of sampling sentences (in No. s)	No of Points generated by probe car data	No. of segments extracted	No. of segments after merging	Percentage of reduction in segments after merging
20	413	19	9	52.63
40	413	8	4	50.00
60	413	4	2	50.00
80	413	3	2	33.33

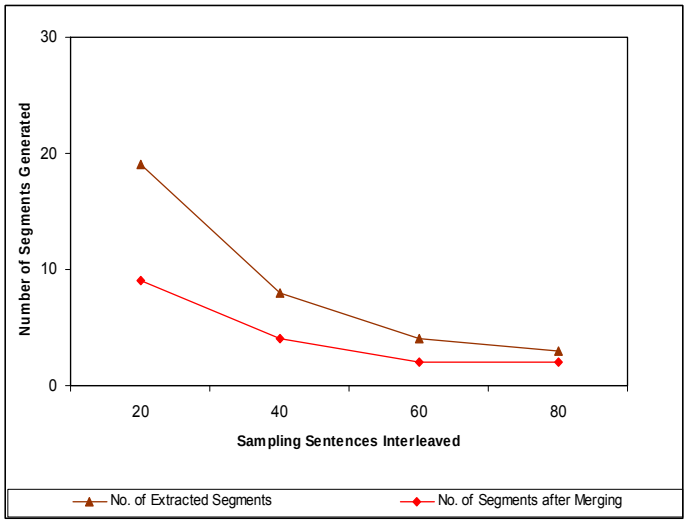


Figure 17. Number of segments generated at different sampling intervals

The accuracy of the map is inversely proportional to the radius of the imaginary circle which is used to draw the road, negligible slope and interval of sampling sentences. Desired accuracy can be obtained by adjusting these values according to the need.

VI. CONCLUSION AND FUTURE WORK

This paper introduces a customized digital road map building process which can be used to build digital maps of desired routes. By changing the input parameters, the accuracy of the route can be altered to required level and route maps for different types of roads, for instance, national highways, state highways, city roads and streets can be drawn. The results of the map building process discussed in this paper are compared with other maps generated by other commercial making software and it is found that the proposed process is simple, yet powerful. This map making process has eliminated the complexity of the previous works carried out in this area and customized road map can be built for commercial and other applications like vehicle tracking, navigation and route guidance systems.

This process can be used to create maps but the process depends on a probe car for data collection. In future, this drawback can be eliminated by collecting GPS data from the GPS enabled vehicles already connected with commercial applications. This work can also be extended to handle roads with multiple lanes.

REFERENCES

[1] Stefen Schroedl, Kiri Wagstaff, Seth Rogers, Pat Langley and Christopher Wilson., "Mining GPS traces for Map Refinement," Data Mining and Knowledge Discovery, vol. 9, pp. 59-87, 2004.

[2] D. Lee, "The Map-Building and Exploration Strategies of Simple Sonar-Equipped Mobile Robot," Cambridge Univ. Press., Cambridge, 1996.

[3] P. Weckesser, and R. Dillmann, "Modeling unknown environment with a mobile robot,", Robotics Autonomous Systems, vol. 23, pp. 293–300, 1998.

[4] Y. L. Ip, A.B. Rad, K.M. Chow and Y.K. Wong, "Segment-based map building using enhanced adaptive fuzzy clustering algorithm for mobile robot applications," Journal of intelligent and robotic systems, vol. 35, pp. 221-245, 2002.

[5] LÜ WeiFeng1, ZHU TongYu, WU DongDong, DAI Hong and HUANG Jian, "A heuristic path-estimating algorithm for large-scale real-time traffic information calculating," Science in China Series E: Technological Sciences, vol. 51, pp. 165-174, Apr. 2008.

[6] B.W. Parkinson, and J.J. Spilker, "Global Positioning System: Theory and Applications," American Institute of Aeronautics and Astronautics, Washington, 1996.

[7] Interface control document. Navstar GPS Space Segment (Navigation User Interfaces), 2000.

[8] Asoke K Talukder and Roopa R Yavagal. "Mobile Computing-Technology, Applications and Service Creation," Tata McGraw Hill Publishing Company, 2005.

[9] X. W. Dai, M.A. Ferman, "A simulation evaluation of a real-time traffic information system using probe vehicles," IEEE ITSC, vol. 1, pp. 12-15, 2003.

[10] R. Kuehne, R.P. Schaefer and J. Mikat, "New approaches for traffic management in metropolitan areas," IFAC CTS, 2003.

[11] T.W. Cheng, D.B. Goldgof, and L.O. Hall, "Fast fuzzy clustering," Fuzzy Sets Systems, vol. 93, No. pp. 49–56, 1998.

[12] R.N. Dave, "Characterization and detection of noise in clustering," Pattern Recognition Lett. vol. 12, pp. 657–664, 1991.

[13] R.N. Dave, "Use of adaptive fuzzy clustering algorithm to detect lines in digital images," Intelligent Robots Computer Vision, vol. VIII, pp. 600–611, 1989.

[14] National Marine Electronic Association, http://www.nmea.org, accessed on 20.04.2010.

AUTHORS PROFILE

G. Rajendran is a second-year Doctoral Research Scholar in the Research and Development Centre of Bharathiar University. He received his Masters degree in Computer Applications and M.Phil degree in Computer Science from Bharathiar University. He passed the National Eligibility Test for lectureship conducted by UGC, the apex body of higher education in India. He is also working as an Assistant Professor of Computer Science at Thiruvalluvar Government Arts College, Rasipuram, India. His research interests include Mobile Computing, Data Mining and programming-language support for massive-scale distributed systems.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Dr. M. Arthanari holds a Ph.D. in Mathematics from Madras University as well as Masters Degree in Computer Science from BITS, Pilani. He was the professor and Head of Computer Science and IT Department at Tejaa Shakthi Institute of Technology for Women, Coimbatore, India. At present he is the Director, Bharathidhasan School of Computer Applications, Ellispettai, Erode, Tamilnadu. He holds a patent issued by the Govt. of India for his invention in the field of Computer Science. He has directed teams of Ph.D. researchers and industry experts for developing patentable products. He teaches strategy, project management, creative problem solving, innovation and integrated new product development for last 36 years.

M. Sivakumar has 10+ years of experience in the software industry including Oracle Corporation. He received his Bachelor degree in Physics and Masters in Computer Applications from the Bharathiar University, India. He is currently doing his doctoral research in Anna University, Coimbatore. He holds a patent for his invention in embedded technology. He is technically certified by various professional bodies like ITIL, IBM Rational Clearcase Administrator, OCP - Oracle Certified Professional 10G, PRINCE2 and ISTQB. His research interests include Embedded Technology, Ubiquitous Computing and Mobile Computing.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Robust stability check of fractional control law applied to a LEO (Low Earth Orbit) Satellite

Ouadiâ EL Figuigui [1], Noureddine Elalami[1]

[1] *Laboratoire d'Automatique et Informatique Industrielle*
EMI, Morocco
(elalami@emi.ac.ma, elfiguigui@gmail.com)

Abstract: **The use of the differentiation and integration of fractional order or non-integer order in systems control is gaining more and more interests from the systems control community. In this paper we will briefly describe the LEO (Low Earth Orbit) satellite systems and recall the theoretical aspects of robust stability check procedure. This procedure will be applied to a LEO satellite that is under the fractional control law. Numerical examples will be analyzed and presented at the end of this document**

Keywords: fractional control, LEO satellite, robust stability

I. . INTRODUCTION

Recently, a lot attention was given to the problem of fractional calculus. There were several works in this area [11],[12],[13],[23]..etc and the author in [1] is presenting for the very first time the robust stability checking procedure for uncertain fractional order linear time invariant (FO-LTI) systems with interval coefficients described in state form. The application of such procedure to LEO satellite was new idea.

The role of an attitude control system for an Earth-pointing spacecraft is to maintain the local-vertical/local-horizontal (LV/LH) attitude with the presence of different environmental disturbances. Most of the time in order to ensure a precise pointing, the satellite requires reaction wheel system to counteract the attitude drifts caused by those perturbations, especially the seculars ones, like torques caused by passive gravity gradient, aerodynamic and solar forces. The reaction wheels are governed by control laws which dictate the amount of torques required to eliminate the drift caused by external factors [9].

In the attitude control design, different approaches have been used, i.e. Proportional Integral Derivative (PID) [18], [25], LQR [17], pole placement techniques [14], etc. All those methods, expressed in different attitude error terminologies are using, very often, Euler angles for small attitude commands while for large attitude maneuvers it makes use of quaternion [7] and direction cosine errors [18].

In [9], the author introduce a fractional control law aiming to stabilize the attitude movement of an earth pointing satellite under the effect of external disturbances, using 3-axis reaction wheels as actuators. The dynamics of the satellite is described by a quasi-bilinear multivariable coupled system. In this study, we apply the robust stability check procedure to the fractional control law system presented in [9].

This paper is organized as follows: In the next section, the general nonlinear equations model of an Earth-pointing satellite attitude dynamics is developed. Then, in section III, the attitude equations are linearized with the nadir attitude position as the origin. This leads to a quasi_bilinear multivariable coupled system. Then, for small maneuvers, the quasi_bilinear term is neglected in order to obtain a linear system. In section IV, we recall some theoretical aspects of robust stability checking procedure. Forward, in section V, we apply this procedure to the LEO satellite system which is subject to fractional control law and we share the related Numerical results. The conclusion is provided in the last section.

II. NONLINEAR MODEL OF THE SATELLITE ATTITUDE DYNAMICS

The attitude motion of the satellite is represented by the Euler equations for the rigid body motion under the influence of external moments, such as the control moment generated by the actuators. Attitude control requires coordinate transformation from LV/LH to The Satellite Coordinate System (SCS) system defined as follows: The LV/LH coordinate system (X_0, Y_0, Z_0) is a right orthogonal system centred in the satellite's centre of mass (SCM). The roll axis, X_0, points along the velocity vector, the pitch axis, Y_0, points in the direction of the negative orbit normal and the yaw axis, Z_0, points in the nadir direction. The SCS system (X_S, Y_S, Z_S) is a right orthogonal system centred in the SCM, parallel to principal moment of inertia axle of satellite. Z_S is parallel to the smallest moment of inertia axis; Y_S is

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

parallel to the largest moment of inertia axis. X_S completes the right orthogonal system.

Consider a satellite with three reaction wheels. The general nonlinear attitude dynamics model can be described as [18], [27] and [21]:

$$\begin{aligned}\bar{I}\dot{\omega}_I^S(t) = &\ \dot{h}_w(t) \ \ \omega_I^S(t)\wedge\bar{I}\omega_I^S(t) \\ &\omega_I^S(t)\wedge h_w(t) + M_g^s(t) + P(t)\end{aligned} \tag{1}$$

where

- $\bar{I}$: Total moment of inertia matrix for the satellite without reaction wheels inertia (3×3).
- $\omega_I^S(t)$: Inertial angular velocity vector in SCS (3×1).
- $h_w(t)$: Angular momentum vector of the wheel cluster(3×1).
- $M_g^s(t)$: Torques due to Earth's gravity gradient (3×1).
- $P(t)$: Disturbance torque due to aerodynamics, solar pressure and other environmental factors. It is assumed to be [18],[21]:

$$\boldsymbol{P}(t) = \begin{bmatrix} 4\times10^{-6} + 2\times10^{-6}\sin(\omega_0 t) \\ 6\times10^{-6} + 3\times10^{-6}\sin(\omega_0 t) \\ 3\times10^{-6} + 3\times10^{-6}\sin(\omega_0 t) \end{bmatrix} \tag{2}$$

where ω_0 is the orbital angular rate.

To keep the satellite attitude earth pointing, the SCS axes must remain aligned with LV/LH axes. The transformation matrix, expressed with Euler Angles (ϕ,θ,ψ), respectively, roll, pitch and yaw angles, is given by [18], [27]:

$$T_{VH/S} = \begin{bmatrix} C_\theta C\phi & C_\theta C_\psi & -S_\theta \\ -C_\phi S_\psi + S_\phi S_\theta C_\psi & C_\phi C_\psi + S_\phi S_\theta S_\psi & S_\phi C_\theta \\ S_\phi S\psi + C_\phi S_\theta C_\psi & -S_\phi C_\psi + C_\phi S_\theta S_\psi & C_\phi C_\theta \end{bmatrix} \tag{3}$$

where: S and C are respectively the sine and the cosine.

The gravity gradient torque $M_g^s(t)$ is given by [18], [5]:

$$\begin{aligned} M_{gx} &= \frac{3}{2}\omega_0^2(I_z - I_y)\sin(2\phi)\cos^2(\theta) \\ M_{gy} &= \frac{3}{2}\omega_0^2(I_z - I_x)\sin(2\theta)\cos(\phi) \\ M_{gz} &= \frac{3}{2}\omega_0^2(I_x - I_y)\sin(2\theta)\sin(\phi) \end{aligned} \tag{4}$$

To describe the satellite kinematics, two important factors are to be taken into account: angular velocity of the body axis frame (SCS) with respect to the reference LV/LH frame $\omega_{VH}^S(p,q,r)$, and the angular velocity of the body frame with respect to inertial axis frame $\omega_I^S(\omega_x,\omega_y,\omega_z)$. These quantities are related to the derivative of the Euler angles as follows [18]:

$$\begin{aligned} p &= \dot{\phi} - \dot{\psi}S_\theta \\ q &= \dot{\theta}C_\phi + \dot{\psi}C_\theta S_\phi \\ r &= \dot{\psi}C_\theta C_\phi - \dot{\theta}S_\phi \end{aligned} \tag{5}$$

and

$$\omega_I^S = \omega_{VH}^S + T_{VH/S}(0 \ \ -\omega_0 \ \ 0)^T$$

III. LINEARIZED EQUATIONS OF MOTION

Assuming small variations of the Eulerian angles (ϕ,θ,ψ), then the transformation matrix becomes:

$$T_{VH/S} = \begin{bmatrix} 1 & \psi & -\theta \\ -\psi & 1 & \phi \\ \theta & -\phi & 1 \end{bmatrix} \tag{6}$$

On the other hand, one obtains that:

$$\dot{\phi} \approx p\ , \dot{\theta} \approx q\ , \dot{\psi} = r \tag{7}$$

and

$$\omega_x = \dot{\phi} - \omega_0\psi,\ \omega_y = \dot{\theta} - \omega_0,\ \omega_z = \dot{\psi} + \omega_0\phi \tag{8}$$

Then the equations of motion (1) and (5) can be linearized about the origin, giving a quasi_bilinear multivariable system:

$$\begin{aligned} \dot{x}(t) = &\ Ax(t) + Bu(t) \\ &+ B\sum_{i=1}^{3}\int_0^t u_i(\xi)d\xi \ \ (C_i x(t)) \\ &+ G\int_0^t u(\xi)d\xi \ + BP(t) \end{aligned} \quad ; x(0) = x_0 \tag{9}$$

where:

- $B\sum_{i=1}^{3}\left(\int_0^t u_i(\xi)d\xi\right)(C_i x(t)) + G\left(\int_i^t u(\xi)d\xi\right)$:Quasi-bilinear term,
- $u(t) = -\dot{h}_w(t)$: Control action,
- $x(t) = (\phi,\theta,\psi,\dot{\phi},\dot{\theta},\dot{\psi})$: State vector,

System matrices A, B, C_i, G defined as

$$\boldsymbol{C}_1 = \begin{bmatrix} 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ -\omega_0&0&0&0&0&-1 \\ 0&0&0&0&1&0 \end{bmatrix} \quad \boldsymbol{C}_2 = \begin{bmatrix} 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ \omega_0&0&0&0&0&1 \\ 0&0&0&0&0&0 \\ 0&0&0&-\omega_0&-1&0 \end{bmatrix}$$

$$\boldsymbol{C}_3 = \begin{bmatrix} 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&-1&0 \\ 0&0&-\omega_0&1&0&0 \\ 0&0&0&0&0&0 \end{bmatrix} \quad \boldsymbol{B} = \begin{bmatrix} 0&0&0 \\ 0&0&0 \\ 0&0&0 \\ 1/I_x&0&0 \\ 0&1/I_y&0 \\ 0&0&1/I_z \end{bmatrix}$$

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

$$A=\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -4\omega_0^2\sigma_1 & 0 & 0 & 0 & 0 & \omega_0(1-\sigma_1) \\ 0 & 3\omega_0^2\sigma_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \omega_0^2\sigma_3 & -\omega_0(1+\sigma_3) & 0 & 0 \end{bmatrix}$$

$$G=\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \omega_0/Ix \\ 0 & 0 & 0 \\ -\omega_0/Iz & 0 & 0 \end{bmatrix}$$

where $\sigma_i = (I_j - I_k)/I_i$ for the (i,j,k) index sets $(1,2,3)$, $(2,3,1)$, and $(3,1,2)$.

Moreover, assuming that the angular velocity components p, q and r are also small, and for slight manoeuvres, one can neglect the quasi bilinear term.

The equation of motion (9) can then be written in the standard form of a linear equations system:

$$\dot{x}(t) = Ax(t) + Bu(t) + BP(t);\ x(0) = x_0 \qquad (10)$$

IV. ROBUST STABILITY CHECK OF FRACTIONAL SYSTEM WITH INTERVAL UNCERTAINTIES - MATHEMATICAL ASPECTS

In this section, we're recalling the definition of fractional system; we're presenting the robust stability checking procedure afterwards.

A. Definition

In this paper, we consider the Riemann-Liouville definition, in which the fractional order integrals are defined as

$$D_a^{-\mu} f(t) = \frac{1}{\Gamma(\mu)} \int_a^t (t-\xi)^{\mu-1} f(\xi) d(\xi) \quad \mu > 0 \qquad (11)$$

While the definition of fractional order derivatives is

$$D_a^{\mu} f(t) = \frac{d}{dt}\left[D_a^{(1-\mu)} f(t)\right] = \frac{1}{\Gamma(1-\mu)} \frac{d}{dt} \int_a^t (t-\xi)^{-\mu} f(\xi) d(\xi) \qquad (12)$$

where $\Gamma(x) = \int_0^{\infty} y^{x-1} e^{-y} dy$ is the Gamma function, $(a,t) \in IR^2$ with $a < t$ and $0 < \mu < 1$ is the order of the operation.

For simplicity we will note $D^{\mu} f(t)$ or $f^{u}(t)$ for $D_0^{\mu} f(t)$

B. Robust Stability Check of Fractional System with Interval Uncertainties [1]:

We consider the following FO-LTI system with interval uncertain:

$$X^{(\alpha)}(t) = \Lambda X(t) + Dw(t) \qquad (13)$$

where:

- α is non integer number;
- $\Lambda \in A^I = [\underline{\Lambda}, \overline{\Lambda}] = [\Lambda^c - \Delta\Lambda, \Lambda^c + \Delta\Lambda]$ with $\Lambda^c = \frac{\Lambda + \Delta\Lambda}{2}$ is a center matrix (normal plant without uncertainties)
- $\Delta\Lambda = \frac{\overline{\Lambda} - \underline{\Lambda}}{2}$ is a radius matrix correspondence interval uncertainties.

The stability condition for system (13) is:

$$\min_i \left|\arg(\lambda_i(\Lambda))\right| > \alpha\frac{\pi}{2} \quad ; \quad i = 1,2,\ldots,N, \quad \forall \Lambda \in A^I$$

In the following subsection, we describe briefly the procedure of checking robust stability using the minimum argument phase. We proceed by introducing two important lemma:

1. Lemma 1 [1]:

Define a sign calculation operator evaluated at Λ^c such as:

$$P^i := sgn\left[\left(u_i^{re} v_i^{re} - u_i^{im} v_i^{im}\right)^T\right] \qquad (14)$$

where $u_i^{re}, v_i^{re}, u_i^{im}$ and v_i^{im} are eigenvectors corresponding to ith eigenvalue of Λ^c. If P^i is constant for all $\Lambda^I, \Lambda^I \in A^I$, then the lower and upper boundaries of the real part of ith interval eigenvalue are calculated as:

$$\underline{\lambda_i^{re}} = \theta_i^{re}\left(\Lambda^c - \Delta\Lambda \circ P^i\right) \qquad (15)$$

where $\theta_i^{re}(\cdot)$ is an operator for selecting the ith real eigenvalue and $C = A \circ B$ are $c_{kj} = a_{kj} b_{kj}$, and as:

$$\overline{\lambda_i^{re}} = \theta_i^{re}\left(\Lambda^c + \Delta\Lambda \circ P^i\right) \qquad (16)$$

1. Lemma 2: [1]

Defining a sign calculation operator evaluated at Λ^c such as:

$$Q^i := sgn\left[\left(u_i^{re} v_i^{im} + u_i^{im} v_i^{re}\right)^T\right] \qquad (17)$$

if Q^i is constant for all $\Lambda^I, \Lambda^I \in A^I$, then the lower and upper boundaries of the imaginary part of ith interval eigenvalue are calculated as:

$$\underline{\lambda_i^{im}} = \theta_i^{im}\left(\Lambda^c - \Delta\Lambda \circ Q^i\right) \qquad (18)$$

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

where $\theta_i^{im}(\cdot)$ is an operator for selecting the ith imaginary eigenvalue, and as:

$$\overline{\lambda_i^{im}} = \theta_i^{im}\left(\Lambda^c + \Delta\Lambda \circ Q^i\right) \tag{19}$$

So, using Lemmas 1 and 2, it is easy to calculate the lower and upper boundaries of interval eigenvalue separately in real part and imaginary part. From above lemma, if P^i and Q^i, $i = 1,\ldots, N$ are calculated, then, interval ranges of eigenvalues are finally calculated as:

$$\lambda_i^I \in \Gamma_i^I := \left\{\left[\underline{\lambda_i^{re}}, \overline{\lambda_i^{re}}\right] + j\left[\underline{\lambda_i^{im}}, \overline{\lambda_i^{im}}\right]\right\} \tag{20}$$

where j represents imaginary part. We define

$$\phi^* = \inf\left(\min_i \left|arg\lambda_i(\Lambda)\right|\right) \qquad i = 1,..N$$

Since the stability condition is given as $\phi^* > \alpha\pi/2$, if we find sufficient condition for this, the stability can be checked. For calculating ϕ^*, the following procedure can be used:

P1. Calculate P_i and Q_i for $i = 1,\ldots, N$.

P2. Calculate $\underline{\lambda_i^{re}}$, $\overline{\lambda_i^{re}}$, $\underline{\lambda_i^{im}}$, and $\overline{\lambda_i^{im}}$ for all $i \in \{1,2,\ldots,N\}$.

P3. Find arguments of phase of four points such as

$$\phi_i^1 = \angle\left(\underline{\lambda_i^{re}}, \underline{\lambda_i^{im}}\right),\ \phi_i^2 = \angle\left(\underline{\lambda_i^{re}}, \overline{\lambda_i^{im}}\right)$$
$$\phi_i^3 = \angle\left(\overline{\lambda_i^{re}}, \overline{\lambda_i^{im}}\right),\ \varphi_i^4 = \angle\left(\overline{\lambda_i^{re}}, \underline{\lambda_i^{im}}\right)$$

in the complex plane.

P4. Find $\phi_i^* = \inf\left\{\left|\phi_i^1\right|, \left|\phi_i^2\right|, \left|\phi_i^3\right|, \left|\phi_i^4\right|\right\}$.

P5. Repeat procedures P3 and P4 for $i = 1,\ldots, N$.

P6. Find $\phi^* = \inf\{\phi_i^*, i = 1,\ldots, N\}$

P7. If $\phi^* > \alpha\pi/2$, then the fractional interval system is robust stable. Otherwise, the stability of system cannot be guaranteed.

V. ROBUST STABILITY OF LEO SATELLITE

The aim of this section is to apply the robust stability checking procedure subject of the section IV. This procedure will be used to proof that our system (10) under the control law presented in [9] is robust stable.

A. Fractional control law

As mentioned above, the LEO satellite atitude dynamics is described, when neglecting the quasi-bilinear term, by the system (10):

$$\dot{x}(t) = Ax(t) + Bu(t) + w(t);\ x(0) = x_0 \tag{21}$$

where $w(t) = BP(t)$ is the perturbation term, and $u(t)$ is the fractional control law applied in order to stabilize the system (10), given by:

$$u(t) = K(x(t) - x_r)^{(\alpha)} \tag{22}$$

where x_r is the attitude reference, equal to zero for nadir pointing.

The linear fractional system is obtained in the form:

$$\dot{x}(t) = Ax(t) - BKx(t)^{(\alpha)} + w(t);\ x(0) = x_0 \tag{23}$$

In the following, only the fractional orders such as $\alpha = 1/p$, $p \in \mathrm{N}^*$ will be considered. Then

$$\dot{x}(t) = Ax(t) - BKx(t)^{(1/p)} + Dw(t) \tag{24}$$

The equation (24) can be written into the following form:

$$X^{(1/p)}(t) = \Lambda X(t) + Dw(t) \tag{25}$$

We note:

$$\begin{aligned}
\left(x^{(1/P)}\right)^{*0}(t) &= x(t),\\
\left(x^{(1/P)}\right)^{*1}(t) &= x^{(1/p)}(t)\\
&\vdots\\
\left(x^{(1/P)}\right)^{*p-1}(t) &= x^{\left(1-\frac{1}{p}\right)}(t)
\end{aligned} \tag{26}$$

and

$$X(t) = \left[\left(x^{(1/p)}\right)^{*0}(t), \left(x^{(1/p)}\right)^{*1}(t), \ldots, \left(x^{(1/p)}\right)^{*p-1}(t)\right]^T \tag{27}$$

$$\Lambda = \begin{pmatrix} 0 & Id & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & Id & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & Id & 0 & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & 0 & 0 & 0 & 0 & Id\\ A & -BK & 0 & 0 & 0 & 0 & 0 \end{pmatrix};\ D = \begin{pmatrix} 0\\ 0\\ 0\\ \vdots\\ \vdots\\ 0\\ Id \end{pmatrix} \tag{28}$$

As mentioned above in the subsection *IV.2*, the system (26) is stable, if and only if:

$$\min_i \left|arg\lambda_i(\Lambda)\right| > \frac{\pi}{2p} \tag{29}$$

The problem of stabilization by state feedback is equivalent to find a matrix K which stabilizes (10); i.e. which checks the stability condition given in (29).

B. Robust stability of the fractional control law

In this section, we suppose that our system is submitted to some perturbations which affect his parameters:

1. Case 1: Perturbation of ω_0 orbital angular rate

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

We consider that, due to the external perturbation, ω_0 varies between $\underline{\omega_0} = \omega_0 - \Delta\omega_0$ and $\overline{\omega_0} = \omega_0 + \Delta\omega_0$. Consequently, *A* varies between $\underline{A}$ and $\overline{A}$. So our system (26) is transformed to FO-LTI system with $\Lambda \in \left[\underline{\Lambda} \quad \overline{\Lambda}\right]$.

For checking the robust stability of system (26), we apply the procedure described in section *IV*.

If $\phi^* > \alpha\pi/2$ the system (26) is robust stable. Otherwise, the stability of system cannot be guaranteed.

- Numerical application

The simulation parameters are the orbital rate $\omega_0 = 0.00104 \quad rad/sec$ and the total moment of inertia matrix for the spacecraft.

$$I = \begin{bmatrix} 4{,}020 & 0 & 0 \\ 0 & 3{,}989 & 0 \\ 0 & 0 & 3{,}010 \end{bmatrix} Kg.m^2$$

The fractional control law which stabilizes the system (10) with these numerical values is $u(t) = -Kx^{\alpha}$ with $\alpha = 0{,}5$ and

$$K = \begin{bmatrix} 0 & 0 & 1 & 83 & 0{,}758 & 3{,}530 \\ 9 & 0 & 23{,}900 & 0 & 2900 & 7 \\ 0 & 0 & 0 & 3{,}200 & 0 & 55 \end{bmatrix}$$

We give below the results of robust stability checking procedure applied to the system (26). These results justifies the robustness of fractional control laws with respect to the variation of ω_0 which can reach $\pm 20\%$.

We note also that the curve of the evolution have a linear behaviour and ϕ^* decrease when $\frac{\Delta\omega_0}{\omega_0}$ increase.

Table I: Results Of Robust Stability Checking Procedure For Variation Of ω_0

$\Delta\omega_0$	$\Delta\omega_0/\omega_0$	$\underline{\omega_0}$	$\overline{\omega_0}$	φ^*
0	0%	0.0010400	0.0010400	0.9334
0.0000416	4%	0.0009984	0.0010816	0.9060
0.0000832	8%	0.0009568	0.0011232	0.8772
0.0001248	12%	0.0009152	0.0011648	0.8476
0.0001664	16%	0.0008736	0.0012064	0.8178
0.0002080	20%	0.0008320	0.0010400	0.7888

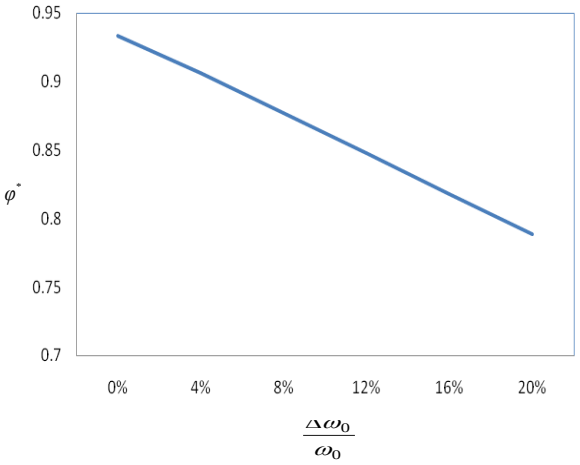


Fig 1: φ^* versus $\frac{\Delta\omega_0}{\omega_0}$

2 .Case 2 : Intrinsic Parameters of LEO Satellite are Perturbed

- Variation of I_x :

We suppose that I_x varied between $I_x - \Delta I_x$ et $I_x + \Delta I_x$ due to external perturbation which modify the form of the satellite.

For the simulation, we consider the same numerical values and fractional control law of subsection *V-2-1*.

The following table resume the results of robust stability checking procedure.

The system is still robust stable until a variation of $\pm 20\%$ of I_x. The ϕ^* decrease when $\frac{\Delta\omega_0}{\omega_0}$ increase.

Table Ii:Results Of Robust Stability Checking Procedure For Variation Of I_x

ΔIx	$\Delta Ix/Ix$	$Ixmin$	$Ixmax$	φ^*
0	0.00%	4.0200	4.0200	0.9334
0.1608	4.00%	3.8592	4.1808	0.8909
0.3216	8.00%	3.6984	4.3416	0.8527
0.4824	12.00%	3.5376	4.5024	0.8244
0.6432	16.00%	3.3768	4.6632	0.8062
0.8040	20.00%	3.2160	4.8240	0.8020

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

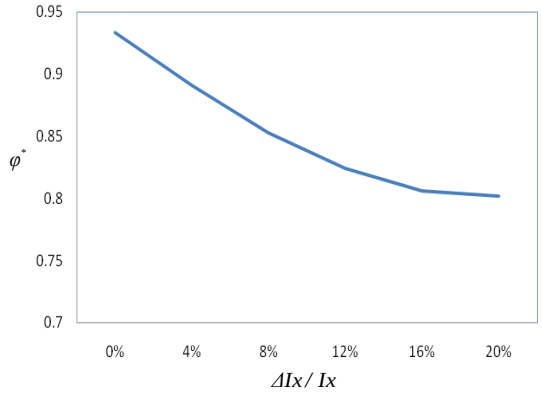


Fig 2: φ^* versus $\Delta Ix/ Ix$

- Variation of I

Using the same values as before, we give bellow the results of robust stability checking procedure.

The system is still robust stable until a variation of $\pm$ 10,53%. The ϕ^* decrease when $\frac{\Delta\omega_0}{\omega_0}$ increase.

Table Iii: Results Of Robust Stability Checking Procedure For Variation Of I

$\Delta I / I$	ϕ^*
0%	0.9334
2.00%	0.8889
4.00%	0.8457
6.00%	0.8165
8.00%	0.8163
10.53%	0.7897

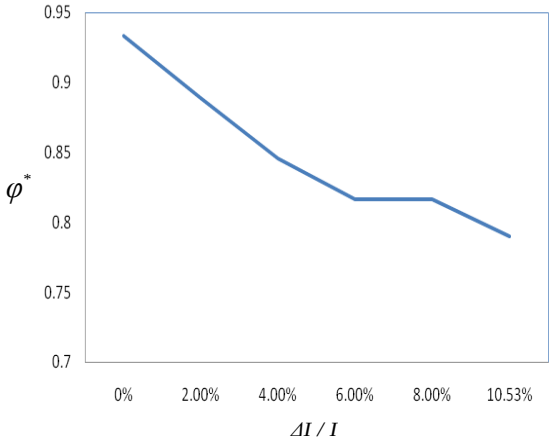


Fig 3: ϕ^* versus $\Delta I / I$

VI. CONCLUSION

In this paper, we assume that the orbital angular rate is subject to some perturbations and, consequently, the LEO satellite system become uncertain. The added value from this work is to prove that the fractional order control present in [9] is robust stable. We used as method the robust stability checking procedure developed in [1]. The mathematical aspects of this procedure were recalled. We have studied also the case of the perturbation of intrinsic parameters of LEO satellite due to the external perturbation which mean that the total moment of inertia is uncertain.

REFERENCES

[1] Y.Q.Chon, H.S.Ahn, I.Podlubny, Robust stability check of fractional order linear time invariant systems with interval uncertainties, signal processing 86 (2006) 2611-2618.

[2] F.A.Devy Vareta, Pseudo-Invariance Sous Groupe de Transformation :un nouveau Concept pour la Commande Robuste, Séminaire Toulousain « Représentation Diffusive et Application »- N°1-nov. 2000.

[3] L. Dorcak, I. Petras, I. Kostial and J. Terpak, State space controller design for the fractional-order regulated system, ICCC 2001, Korynica, Poland, pp. 15-20.

[4] M. M. Dzhebashyan, Integral transforms and representation of functions in the complex plan, Nauka Moscou 1966.

[5] P.C. Hughes, Spacecraft Attitude Dynamics, John Wiley & Sons, USA New York, 1986

[6] B. Kim, E. Velenis, P. Kriengsiri & P. Tsiotras, A Spacecraft Simulator for Research and Education, AAS 01-367.

[7] B.J. Kim, H. Lee and S.D. CHOI, Three-axis Reaction Wheel Attitude Control System for KITSAT-3 Microsatellite, IFAC Conference in Autonomous and Intelligent Control in Aerospace, Beijing, 1995.

[8] A. Kailil, Architecture et Analyse Dynamique par la Méthode des Eléments Finis de la plate forme d'un microsatellite, DESA, CRES, Mohammedia Engineers School, Maroc, 2000.

[9] A. Kailil, N. Mrani, M. Abid, M. Mliha Touati, S. Choukri, N. Elalami, Fractional regulators for spacecraft attitude stabilization, accepted in the 22nd AIAA-ICSSC, Monterey, California 9-12 May 2004.

[10] S.Ladaci,J.J.Loiseau,A.Charef,Stability Analysis of Fractional Adaptive High-Gain Controllers for a class of Linear Systems General case,2006, IEEE.

[11] S.Ladaci,A.Charef, Mit Adaptive Rule with Fractional Integration, in Proceedings CESA 2003 IMACS Multiconference Computational Engineering in Systems Applications, Lille, France.

[12] S.Ladaci, A.Charef, On Fractional Adaptive Control, nonlinear Dynamics, vol.43,n°4, March 2006.

[13] S.Ladaci, A.Charef, An Adaptive Fractional PI D. Controller, in Proceedings TMCE 2006 international Symposium series on Tools and Methods of Competitive Engineering, Ljubljana, Slovenia, April, April18-22.

[14] CH. Lubich, Discretized fractional calculus, SIAM J. Math. Anal. Vol. 17 No. 3 May 1986. P

[15] D. Matignon, Fractals et loi d'échelle, Edition Hermes 2002.

[16] D. Matignon, Stability results for fractional differential equations with applications to control processing, in Computational Engineering in Systems and Application multiconference, vol. 2, pp. 963-968, IMACS, IEEE-SMC. Lille, France, July 1996.

[17] K.L. Musser, W.L. Elbert, Autonomous Spacecraft Attitude Control using Magnetic Torquing only, Proceeding of flight mechanics estimation theory symposium, NASA, 1986, pp. 23-38.

[18] M. J. Sidi, Spacecraft Dynamics and control, Cambridge University Press, Cambridge, UK, 1997.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

[19] S.G. Samko, A. A. Kilbas & O.I. Maritcheu, Fractional integrals and derivative: theory and application, Gorden & Breach, 1987.

[20] A. Skullestad, J. Gilbert, H_∞ Control of a Gravity Gradient Stabilised Satellite, Control Engineering Practice 8 (2000) 975-983.

[21] P. Tsiotras, H. Shen and C. Hall, Satellite Attitude Control and Power Tracking with Momentum Wheels, AAS 99-317.

[22] C. Valentin-Charbonnel, G. Duc and S. Le Ballois, Low-order Robust Attitude Control of an Earth Observation Satellite, Control Engineering Practice 7 (1999) 493-506.

[23] B.M.Vinagre,I.Petras, I.Podlubny and Y.Q.Chen, Using Fractional Order Ajustement Rules and Fractional Order References Models in Model-Reference Adaptive Control, Nonlinear Dynamics, vol.29.

[24] B. Wie, Space Vehicle Dynamics and Control, AIAA Education Series, 1998.

[25] C. Witford, D. Forrest, The CATSAT Attitude Control system, Proceeding of the 12th Annual AIAA/USU Conference on Small Satellite, 1996.

[26] C. H. Won, Comparative study of various methods for attitude control of LEO satellite, Aerospace Science and Technology. 1270-9638. 99/05/ Elsevier, Paris.

[27] J. R.Wertz, Spacecraft Attitude Determination and control, Kluwer Academic Publishers, Dordrecht, Holland, 1978.

[28] K. Zhou, JC. Doyle & R. Glover, Robust and Optimal control, Prentice Hall, 1996

PERFORMANCE EVALUATION OF GENETIC ALGORITHM FOR SOLVING ROUTING PROBLEM IN COMMUNICATION NETWORKS

Ehab Rushdy Mohamed
Faculty of Computer and Informatics
Zagazig University
Zagazig, Egypt
ehab.rushdy@gmail.com

Ibrahim Elsayed Zidan
Faculty of Engineering
Zagazig University
Zagazig, Egypt
ibrahim.zidan.123@gmail.com.

Mahmoud Ibrahim Abdalla
Faculty of Engineering
Zagazig University
Zagazig, Egypt
mabdalla2010@gmail.com

Ibrahim Mahmoud El-Henawy
Faculty of Computer and Informatics
Zagazig University
Zagazig, Egypt
. i.m.elhenawy@gmail.com.

***Abstract* —There has been an explosive growth in both computer and communication networks since the last three decades. Communication networks pervade our everyday life, from telephony system, to airline reservation systems, to electronic mail services, to electronic bulletin boards, to the internet. Routing problem is one of the most important issues facing the development, improvement and performance of communication networks. Recently, there has been increasing interest in applying genetic algorithms to problems related to communication networks. This study evaluates the genetic algorithm used for finding the shortest path in communication network. The paths result from applying genetic algorithm could be used in establishing routing table for network protocols. The genetic approach is thought to be an appropriate choice since it is computationally simple, provide powerful search capability, and has the ability to move around in the solution space without a dependence upon structure or locality.**

The performance of the genetic algorithm is compared to Dijkstra algorithm, which is widely used in most network protocols. This is realized using different simulated networks to clarify the advantages and deficiencies of each algorithm. The relative performance of the two algorithms is judged on the basis of delay and adaptation.

I. INTRODUCTION

One of the most common problems encountered in analysis of networks is the shortest path problem: finding a path between two designated nodes having minimum total length or cost. In many applications, however, several criteria are associated with traversing each edge for a network. For example, cost and time measures are both important in most networks, as are economic and ecological factors. As a result, there has been recent interest in solving the shortest path problem and many algorithms are proposed to solve this problem [1], [2], [3], [4]. Shortest path problem is a classical research topic. It was proposed by Dijkstra in 1959 and has been widely researched [5].

A. Dijkstra Algorithm

The Dijkstra algorithm is considered as the most efficient method. It is based on the Bellman optimization theory. Dijkstra routing algorithm is widely used in many applications because it is very simple and easy to understand. This algorithm finds the shortest paths from a source to all other nodes. To do this it requires global topological knowledge (the list of all nodes in the network and their interconnections, as

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

well as the cost of each link). In most general the weight of each link could be computed as a function of the distance, bandwidth, average traffic, communication cost, mean queue length, average delay, and other factors. In most general the weight of each link could be computed as a function of the distance, bandwidth, average traffic, communication cost, mean queue length, average delay, and other factors [6].

The parameter *D(v)* is considered as the distance (sum of links weights along any path) from source node *1* to node *v*, and the parameter *L(i,j)* is considered as the given cost between node *i* and node *j*. There are then two parts to the algorithm: an initialization step, and a step to be repeated until the algorithm terminates:

1. Initialization

Set $N = \{1\}$. For each node *v* not in Set *N, D(v)* is set to *L(1,v)*. The value ∞ is used for nodes that are not connected to node *1*; any number larger than maximum cost or distance in the network would suffice.

2. At each subsequent step

Find a node *w* not in *N* for which *D(w)* is a minimum and add *w* to *N*. Then *D(v)* is updated for all nodes remaining that are not in *N* by computing $D(v) = Min [D(v) , D(w)+L(w,v)]$

Step 2 is repeated until all nodes are in N.

The Dijkstra algorithm is widely used in most popular routing protocols because of its simplicity and efficiency. But when the network is very big, then it becomes inefficient since a lot of computations need to be repeated. The efficient set of paths may be very large, possibly exponential in size. Thus the computational effort required for solving the problem can increase exponentially with the problem size [7].

B. Genetic Algorithm Mechanism

The genetic approach is a special kind of stochastic search algorithm, it based on the concept of natural selection and genetics. It provides a powerful search capability and has the ability to move around in the solution space without the structure and locality. This approach identifies solutions that are closest to the ideal solution as determined by some measures of distance. Genetic algorithm constitutes the increasingly large part of evolutionary calculation techniques, which form the artificial intelligence .As it's obvious in its name, genetic algorithm, forming evolutionary popular technique, inspired from evolution theory of Darwin. Any kind of problems which involves genetic algorithm is solved through the application of artificial evolution technique. Genetic algorithm is used to solve problems that are hard to be solved by applying conventional methods. In general terms, genetic algorithm has three field of application. They are; optimization, practical industrial applications, and categorization systems [8], [9].

Genetic Algorithm starts with a set of solution, which is identified with chromosomes and known as population. Resolutions that have come out from a population are applied to the next one with the expectation of positive improvements. The selected group is used for creation of a new population according to their compatibility. Nevertheless, it's likely that the compatible ones produce better solutions. This would be continued until the expected solutions are obtained. A simple genetic algorithm consists of the following steps:

1. Initialization: a random population of *n* chromosomes (appropriate solution of the problem) is created.

2. Fitness: each *x* chromosome is evaluated using the fitness f (x).

3. New population: a new population is created; this is done by repeating the following steps until the new population is complete

a. Selection: two parental chromosomes are selected from the population according to their fitness.

b. Crossover: a new member is created; parents are cross-fertilized according to possibility of crossover. If cross-fertilization is not applied, new member is a copy of a mother or father.

c. Mutation: the place of the new member is changed according to the possibility of fission.

d. Addition: the new member is added to the new population.

4. Alteration: new generated population is used when algorithm is re-applied.

5. Test: If the result is convincing, algorithm is concluded and the last population is presented as the solution.

6. Cycle: Return to the second step.

As seen above, the structure of the genetic algorithm is quite general and can be applied to any kind of problem. Identification of chromosomes is generally done by the numbers in double set. Members that are used for the crosswise must be selected among the best ones. There is no completion criterion of GA. Having a satisfactory result or guaranteeing the convergence could be used as criteria for completion of the algorithm.

The most important parts of GA are the processes of crossover and mutation. These processes are started with a unit of probability, and in most of the cases

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

applied randomly. This helps to get satisfactory results. A chromosome should include information on solution that it represents. Each chromosome is set up with binary series. Each number that named bit in this series can represent a characteristic of the solution. Or, a serial, on its own, would indicate a number. Expressing the chromosomes with the set of numbers in the binary series is the most common representation form; however, also integer and real numbers can be used [9]. The reason for selecting of binary series is as the following: first of all it is simple, and secondly, it is processed by the computer easier and faster. The reproduction process is a process which is applied according to certain selection criteria to reproduce new generation. A selection criterion takes the compatibility as a basis and selects the compatible members. At later stage, it is possible that more compatible new members emerge from those members that are subjected to crosswise and fission. All members may be selected in terms of compatibility or some are selected randomly and transferred to next generation. Crossover can be applied after the decision for representation of chromosomes taken. Crossover is a process which is applied through the deduction of some genes from parents to create new members. There is a need of selection of individuals to constitute parents. According to theory the fittest individuals must be survive to leave descendants. This selection can be based on several criteria. Examples are Roulette selection, Boltzmann selection, tournament selection, sorted selection [8].

Genetic algorithms, as powerful and broadly applicable stochastic search and optimization techniques, are the most widely known types of evolutionary computation methods today. Recently, the genetic algorithm community has turned much of its attention to optimization problems in the field of communication and computer networks, resulting in a fresh body of research and applications [10], [11], [12], [13], [14], [15], [16].

II. Genetic Approach for Solving the Shortest Path

Routing problem is one of the most important issues facing the development, improvement and performance of communication networks. Many ideas and methods have been proposed to solve routing problems [17], [18], [19], [20], [21]. One of the most common problems encountered in networks is the shortest path problem. The shortest path problem is defined as that of finding a minimum length or cost path between a given pair of nodes. Shortest path problem is a classical research topic.

A. Problem Formulation

An undirected graph $G = (V,E)$ comprises a set of nodes $V=\{1,2,\ldots\ldots,n\}$ and a set of edges $E \in V \times V$ connecting nodes in V. For each edge in the graph, there is a nonnegative number c_{ij} represents the cost, distance, and others of interest, from node i to node j. A path from node i to node j is a sequence of edges $(i,l),(l,m),\ldots,(k,j)$ from E in which no node appears more than once. A path can also be represented using a sequence of nodes $(i,l,m\ldots k,j)$. x_{ij} is an indicator variable. It is equal 1 if the edge (i,j) is included in the path and zero otherwise.

The shortest-path problem is formulated as follows:

$$minimize \quad f(x) = \sum\sum c_{ij} x_{ij} \tag{1}$$

$$subject\ to \quad \sum_j x_{ij} \le 2, \qquad \forall i \in V \tag{2}$$

$$\sum_{j \ne k} x_{ij} \ge x_{ik}, \quad \forall (i,k) \in E, \ \forall i \in V \backslash \{1,n\} \tag{3}$$

$$\sum_j x_{1j} = \sum_j x_{jn} = 1, \qquad \forall (i,j) \in V \tag{4}$$

where constraints (2) and (3) together imply that any node other than nodes 1and node n has either 0 or 2 nonzero incident edges. Constrain (4) makes node 1and n the endpoints of the path [9].

B. Encoding The Path Of The Graph

The difficult part of developing a genetic algorithm for this problem is how to encode a path in graph into a chromosome. A priority based encoding method can potentially represent all possible paths in a graph. Besides, the capability of representing a path for undirected graph is one of the most significant features of the priority based encoding method. This difficulty of finding a representation for the paths for a graph in genetic approach arises because of the existence of variable number of nodes in the paths, and the available number is $n-1$ for an n-order graph. Also, a path consists of consecutive prepared number of edges. A random sequence of edges doesn't correspond to a path. These problems were solved by encoding some guiding information for establishing a path in a chromosome but not a path itself.

Any chromosome is defined by two factors. The first is locus that represents the position of gene in the structure, while the second is allele represents the value

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

of gene. Simply, the node is represented by locus and the priority of representing the node in constructing a path with respect to other ones.

The mapping between encoding and path is many-to-one, which means that different chromosomes may produce an identical shortest path. But it is easy to prove that the probability of occurrence of many-to-one mapping is very low [8]. Therefore, in most cases, there are no trivial genetic operations associated with the encoding. It is also easy to verify that any permutation of the encoding corresponds to a path, so that most existing genetic operators can easily be applied to the encoding. Also, any path has a corresponding encoding; therefore, any point in solution space is accessible for genetic search.

A path growth procedure is considered to generate a path from chromosomes. The procedure generates a path from initial node to end node by appending eligible edges into the path consecutively. At each step, several edges are considered. The edge added to a partial path is always the edge incident to the node with the highest priority, which extends the partial path from the terminal node [9].

C. Genetic Algorithm Procedure

Chromosomes are evaluated using a measure of fitness during each iteration of a genetic algorithm. There are three major steps included in the evaluation:

1. Convert the chromosome to a path.
2. Calculate the objective values.
3. Convert the objective values to fitness values.

The roulette wheel approach, a type of fitness-proportional selection, was adopted as the selection procedure. The elitist method was combined with this approach to preserve the best chromosome in the next generation and overcome stochastic errors of sampling. With the elitist selection, if the best individual in the current generation is not reproduced in the new generation, one individual is removed randomly from the new population and the best one is added to the new population [7], [8].

Genetic approach uses the comprise approach based fitness assignment in solving the shortest path problem. The compromise approach is regarded as a type of mathematical formulation of goal-seeking behavior in terms of a distance function. The compromise approach identifies solutions that are closest to optimum. The overall genetic approach is illustrated as follows:

- Step 1: The data are entered and genetic parameters are set.
- Step 2: The initial population is generated randomly.
- Step 3: The paths are encoded into chromosomes.
- Step 4: The parent chromosomes are selected
- Step 5: The objectives are calculated and evaluated for each individual.
- Step 6: The next generation is produced by applying crossover and mutation to the parent chromosomes.
- Step 7: The new generations replaced the current ones.
- Step 8: If the maximum generation is reached then procedure is stopped; otherwise; go to step 4.

III. SIMULATION AND RESULTS

Simulation experiments are carried out for a number of networks with different size. The objective is to investigate and evaluate the performance of the genetic algorithm and Dijkstra algorithm as well using different simulated networks .The sizes of simulated networks considered are 30, 50, 100, 120 and 150 nodes. Assumptions for cost links are considered for the simulated networks. The topology of 30-node simulated network used is illustrated in Figure 1.

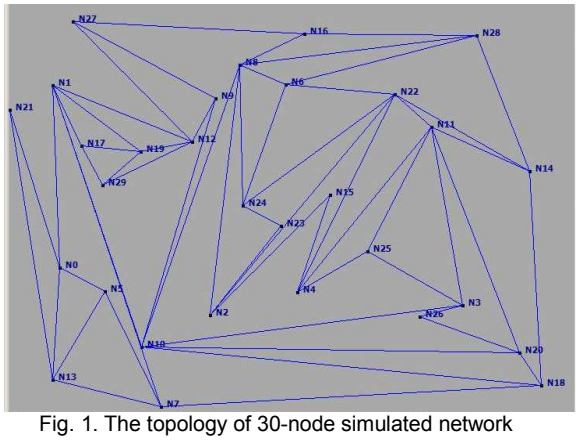


Fig. 1. The topology of 30-node simulated network

It should be pointed out that the performance of the genetic algorithm depends on the choice of the parameters like population size, crossover rate and mutation rate. In this simulation, various values of these parameters are experimented. The optimal values are found to be 800, 0.6 and 0.1 for population size, cross over rate and mutation rate respectively as illustrated in Table I.

TABLE I. THE INITIAL VALUES FOR GENETIC ALGORITHM

Parameter	**Value**
Population size	800
Mutation rate	0.6
Cross over rate	0.1
Number of generation	100

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

As an example, Table II shows the source and destination for each pair of nodes in the 30-node simulated network to be applied on both genetic algorithm and Dijkstra algorithm.

TABLE II. THE FOUR SOURCE-DESTINATION PAIRS FOR 30-NODE SIMULATED NETWORK

Case	Source node	Destination node
1	N11	N12
2	N8	N25
3	N6	N29
4	N0	N22

Table III shows the results of implementing the genetic algorithm for the first route in the 30-node simulated network. Table IV shows the results of implementing genetic algorithm for the second route in the 30-node simulated network. The results of implementing genetic algorithm for the third route in the 30-node simulated network is illustrated in Table V. Table VI shows the results of implementing genetic algorithm for the fourth route in the 30-node simulated network. Figure 2 illustrates the variation of cost with respect to number of generation for the four routes in the 30-node simulated network.

Table VII illustrates the results of applying Dijkstra algorithm on the four pairs of nodes in the 30-node simulated network.

TABLE III. THE RESULTS OF APPLYING THE GENETIC ALGORITHM ON THE 1ST ROUTE IN 30-NODE SIMULATED NETWORK

Path	Cost	Running time in sec	Generation number
N11-N20-N18-N7-N1-N10-N9-N12	114.82	4.449	100
N11-N22-N14-N18-N10-N1-N12	91.30	10.556	200
N11-N22-N4-N25-N3-N10-N1-N12	88.66	14.008	250
N11-N22-N2N-N8-N16-N27-N12	77.46	19.892	400
N11-N3-N10-N9-N12	59.18	29.327	600
N11-N3-N10-N9-N12	59.18	34.563	700
N11-N3-N10-N9-N12	59.18	39.437	800

TABLE IV. THE RESULTS OF APPLYING THE GENETIC ALGORITHM ON THE 2 ND ROUTE IN 30-NODE SIMULATED NETWORK

Path	Cost	Running time in sec	Generation number
N8-N6-N24-N23-N2-N22-N4-N25	67.16	10.034	200
N8-N24-N6-N28-N14-N11-N25	62.39	15.786	300
N8-N28-N14-N11-N4-N25	57.70	21.983	400
N8-N2-N15-N4-N25	43.70	25.556	500
N8-N6-N22-N11-N25	25.54	34.984	700
N8-N6-N22-N11-N25	25.54	45.658	900
N8-N6-N22-N11-N25	25.54	48.772	1000

TABLE V. THE RESULTS OF APPLYING THE GENETIC ALGORITHM ON THE 3RD ROUTE IN 30-NODE SIMULATED NETWORK

Path	Cost	Running time in sec	Generation number
N6-N8-N28-N16-N27-N12-N19-N29	71.71	10.656	200
N6-N8-N16-N27-N9-N12-N19-N17-N29	53.43	14.331	300
N6-N8-N16-N27-N12-N19-N17-N29	50.26	19.711	400
N6-N8-N16-N27-N9-N12-N29	49.24	25.000	500
N6-N8-N28-N16-N27-N12-N29	46.07	29.886	600
N6-N8-N28-N16-N27-N12-N29	46.07	33.223	700
N6-N8-N28-N16-N27-N12-N29	46.07	38.119	800

TABLE VI. THE RESULTS OF APPLYING THE GENETIC ALGORITHM ON THE 4TH ROUTE IN 30-NODE SIMULATED NETWORK

Path	Cost	Running time in sec	Generation number
N0-N21-N13-N7-N18-N14-N28-N6-N22	116.67	15.980	300
N0-N1-N19-N29-N12-N27-N16-N28-N14-N22	96.65	22.879	400
N0-N13-N5-N7-N18-N14-N11-N22	79.94	27.870	500
N0-N13-N7-N18-N14-N11-N22	71.63	40.676	800
N0-N1-N10-N8-N6-N22	66.06	51.922	1000
N0-N1-N10-N8-N6-N22	66.06	54.398	1100
N0-N1-N10-N8-N6-N22	79.94	60.933	1200

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

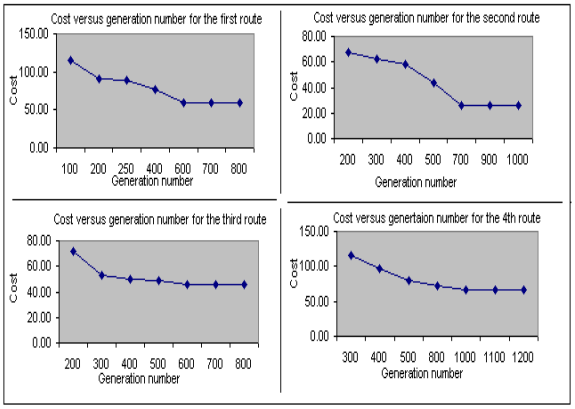


Fig. 2. Variation of cost with respect to number of generation for 30-node simulated network

TABLE VII. THE RESULTS OF APPLYING DIJKSTRA ALGORITHM ON THE 4TH ROUTE IN 30-NODE SIMULATED NETWORK

Case	Path	Cost	Time in sec
1	N11-N3-N10-N9-N12	59.18	0.125
2	N8-N6-N22-N11-N25	25.54	0.058
3	N6-N8-N28-N16-N27-N12-N29	46.07	0.086
4	N0-N1-N10-N8-N6-N22	66.06	0.172

It is pointed out that the genetic algorithm can find the optimum path that is exactly obtained by applying Dijkstra algorithm. Furthermore, various candidate paths close to optimum are obtained using the genetic algorithm. Also, the associated time to get various paths and generation number are illustrated in these tables. It is noted that increasing the generation number leads to get paths more close to optimum as shown in Figure 2. Also, paths more close to optimum are consumed more time than others.

The same results are achieved using simulated networks with size of 50, 100, 120 and 150 nodes.

The only concern during implementing the genetic algorithm is that genetic algorithm needs much running time compared to Dijkstra algorithm, this is illustrated in Table VIII. It is noticed that Dijkstra algorithm is working efficiently and it is implemented in the permitted time with network with size till 150 nodes.

Also, the relation between the average running time of the genetic algorithm and the network size is illustrated by applying the genetic algorithm on different simulated networks of 30, 50, 100, 120 and 150 nodes. It is noticed from the results shown in Figure 3 that the running time needed to implement the genetic algorithm is increased as result of increasing the network size. On the other side, Dijkstra algorithm is not affected with the increase of the network size.

TABLE VIII. COMPARISON BETWEEN THE GENETIC ALGORITHM AND DIJKSTRA ALGORITHM IN ACCORDANCE TO RUNNING TIME

Simulated Network	Average running time in sec	
	Genetic algorithm	Dijkstra algorithm
30-node network	36.530	0.126
50-node network	43.830	0.093
100-node network	48.440	0.112
120-node network	55.135	0.137
150-node network	69.055	0.106

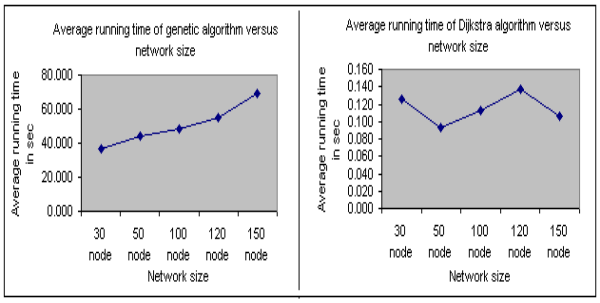


Fig. 3. The average running time of the genetic algorithm and Dijkstra algorithm using five simulated networks with different sizes

Therefore, the genetic algorithm is considered to be a promising algorithm that produces multiple paths close to optimum in addition to the optimum path itself and also can combine with other algorithm to form a hybrid technique that can be used in multiple applications and environments.

IV. CONCLUSION

In this study, the analysis and performance evaluation of the genetic algorithm that used to solve the shortest path problem are presented. The genetic algorithm is experimented using different simulated networks with sizes of 30, 50, 80, 100,120 and 150 nodes. A comparison between both the genetic algorithm and Dijkstra algorithm is illustrated .It should be pointed out that the performance of the genetic algorithm depends on the choice of the parameters like population size, crossover rate and mutation rate. In the simulation, various values of these parameters are experimented before selecting the ones that achieve the desired results. The desired values of population size, crossover rate and mutation rate parameters are 800, 0.6

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

and 0.1 respectively. From the testing and results analysis, the following results are achieved:

1. Genetic algorithm is able to find the optimum solution achieved by Dijkstra algorithm.
2. Increasing the number of generation leads to obtaining paths more close to optimum
3. Genetic algorithm achieves the desired results with much running time comparing to Dijkstra algorithm
4. The running time of implementing the genetic algorithm is increased as a result of increasing the network size.
5. Dijkstra algorithm is working efficiently and it is implemented in the permitted time with network with size till 150 nodes
6. Genetic algorithm is able to find alternative paths close to the shortest path; consequently, these results can be used in establishing routing table for nodes. These alternative paths can be generated using different genetic operators, which can be invoked at a specific probability. As a result, the load and utilization of the paths in communication networks will be slightly reduced and load balance can be achieved using multiple paths close to optimum to route traffic between two nodes.

REFERENCES

[1] D. Awduche, A. Chiu, A. Elwalid, I. Widjaja, and X. Xiao, "Overview and principles of Internet traffic engineering", Internet RFC 3272, May 2002.

[2] [2]A. Montes, "Network shortest path application for optimum track ship routing", California, 2005.

[3] S. F. Wu, F. Y. Wang, Y. F. Jou, and F. Gong, "Intrusion detection for link-state routing protocols", In IEEE Symposium on Security and Privacy, 1997.

[4] M. T. Goodrich, "Efficient and Secure Network Routing Algorithms", Johns Hopkins University, 2001.

[5] E. Dijkstra, "A note on two problems in connection of graph", Numerical Mathematical, 1:269-271, 1959.

[6] M. Pioro, D.Medhe, Routing, Flow, and Capacity Design in Communication and Computer Networks, Morgan Kauffman series, 2004.

[7] Y. Li, R. He, Y. Guo, "Faster Genetic Algorithm for Network Paths", The Sixth International Symposium on Operations Research and Its Applications (ISORA'06), China, 2006.

[8] M. Gen, R. Cheng, Genetic Algorithms & Engineering Optimization, Wiley Series in Engineering and Automation, 2000.

[9] D. Goldberg, Genetic Algorithms in Search, Optimization, and Machine learning, Addison Wesley, 1989.

[10] P. Sateesh, S. Ramachandram," Genetic Zone Routing Protocol" Journal of Theoretical and Applied Information Technology, 2008.

[11] M. R. Masillamani, A. V. Suriyakumar, G. V. Uma, "Genetic Algorithm for Distance Vector Routing Technique", AIML 06 International Conference, Sharm EL Sheikh, Egypt, 2006.

[12] N. Selvanathan and W. Jing, "A genetic algorithm solution to solve the shortest path problem in OSPF and MPLS", Malaysian Journal of Computer Science, Vol.16 No. 1, 58-76, 2003.

[13] J. Cunha, "Map algorithm in routing problems using genetic algorithm, IFORS Triennial Conference – Edinburgh/Scotland, 2002.

[14] B. Fortz, M. Thorup, "Internet traffic Engineering by Optimization OSPF Weights", Proc. IEEE Infocom, 2000.

[15] M. Ericsson, M.G.C. Resende, P.M. Pardalos, " A Genetic Algorithm for the weight setting problem in OSPF routing", Journal of Combinatorial Optimization, 2001.

[16] O. Akbulut, O. Osman, O. N. Ucan, "Computer Network Optimization Using Genetic Algorithm", The journal of Electrical & Electronics Engineering, Istanbul University, Vol. 6, 245-250, 2006.

[17] D. Staehle, S. Koehler, U. Kohlhhass, "Towards an optimization of IP Routing by Link Cost Specification ", University of Wurezburg, 2000.

[18] J. Milbrant, S. Koehler, D. Staehle, L. Berry, " Decomposition of Large IP Networks for Routing Optimization", Technical Report 293, University of Wurezburg, 2002.

[19] G. L. Li and P. W. Dowd , " An analysis of network performance degration induced by workload fluctuations", IEEE/ACM Trans. on networking, 3, No. 4, 433-440,1995.

[20] F. Xiang, L. Junzhou, W. Jieyi, G. Guanqun, " QoS routing based on genetic algorithm", Computer Communication 22, 1392-1399, 1999.

[21] K. Vijayalakshmi, S. Radhakrishnan, " Dynamic Routing to Multiple destinations in IP Networks using Hybrid Genetic Algorithm" International Journal of Information Technology, 2008.

Testing Equivalence of Regular Expressions

Keehang Kwon
Department of Computer Engineering
Dong-A University
Busan, Republic of Korea
khkwon@dau.ac.kr

Hong Pyo Ha
Department of Computer Engineering
Dong-A University
Busan, Republic of Korea
hompoyo@hotmail.com

Abstract— We propose an algorithm that tests equivalence of two regular expressions. This algorithm is written in the style of a sequent proof system. The advantage of this algorithm over traditional algorithms is that it directly captures the real essences regarding language equivalence. As a consequence, our algorithm extends easily to other larger languages with variables**.**

Keywords- *regular expression, equivalence, language, algorithm***.**

1.INTORDUCTION

Regular expressions [1] have gained much interest for applications such as text search or compiler components. One central test related to regular expressions are to test, given two regular expressions, whether two regular expressions are equivalent. The equivalence of regular expressions is useful for simplifying regular expressions. For example, 0 + 01* can be simplified to 01*. Unfortunately, the traditional algorithms[4] focus too much on the equivalence of regular expressions and finite automatas. As a consequence, the resulting algorithm is not so intuitive, as they convert regular expressions to finite automatas for testing equivalence. .
Furthermore, this technique does not extend well to other grammars such as context-free grammars such as context-free grammars. .
This paper introduces an algorithm for testing equivalence of regular expressions. It is simple, easy to understand, nondeterministic and some resemblance to the proof theory of intuitionistic linear logic [2]. .
In addition, it is a simple matter to observe that this algorithm extends well to more general grammars. Our algorithm thus captures the essence of language equivalence. In this paper we present our algorithm,show some examples of its workings, and discuss further improvements. The remainder of this paper. The remainder of this paper is structured as follows. We describe our algorithm in the next section. In Section 3, we present some example. Section 4 concludes the paper with some considerations for further improvements.

2.THE LANGUAGE

The Φ –free regular expression is described by *G*-formulas given by the syntax rules below:

$$G ::= \varepsilon \mid a \mid G \bullet G \mid G+G \mid G^*$$

In the rules above, a represents the set $\{a\}$. ε represents $\{ \varepsilon \}$. Φ represents the empty set. $F \bullet G$ represents the concatenation of two sets F and G. The Kleene closure of *G*-*G**- indicates there are many number of *G*, *i.e.*, $G \bullet \ldots \bullet G$. We write GG in place of $G \bullet G$.
The regular expressions have a number of laws for their equivalence.

For instance, Φ is the identity for union: $\Phi + L$ and $L + \Phi = L$ where L is any regular expression.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Similarly, ε is the identity for concatenation: $\varepsilon L = L\varepsilon = L$.

On the other hand, Φ is the annihilator for concatenation : $\Phi L = L\Phi = \Phi$.

The question of whether two regular expressions are equivalent is quite complex. We will present an algorithm for this task in the style of a proof system. Let G be a regular expression and $\Gamma_1,\ldots,\Gamma_n$ be a list of regular expressions. Then $\Gamma_1,\ldots,\Gamma_n \vdash G$ – the notion that G is subset of the concatenation of $\Gamma_1,\ldots,\Gamma_n$ – is defined as follows :

Algorithm for Subset Relation

$$\frac{\Gamma \vdash \Delta}{\epsilon,\Gamma \vdash \Delta} eL \qquad \frac{}{\vdash \epsilon} \epsilon R$$

$$\frac{\Gamma,\rho,\psi \vdash \Delta}{\Gamma,\rho \bullet \psi \vdash \Delta} \bullet L \quad \frac{\Gamma_1 \vdash \rho \; \Gamma_2 \vdash \psi}{\Gamma_1,\Gamma_2 \vdash \rho \bullet \psi} \bullet R$$

$$\frac{}{\rho \vdash \rho} Axiom \qquad \frac{\Gamma \vdash \rho \; \Gamma \vdash \psi}{\Gamma \vdash \rho + \psi} + R$$

$$\frac{\Gamma,\rho \vdash \psi}{\Gamma,\rho+\psi \vdash \Delta} + L_1 \quad \frac{\Gamma,\psi \vdash \Delta}{\Gamma,\rho+\psi \vdash \Delta} + L_2$$

$$\frac{\Gamma,\rho^*,\rho^* \vdash \Delta}{\Gamma,\rho^* \vdash \Delta} CL \quad \frac{\Gamma,\rho \vdash \Delta}{\Gamma,\rho^* \vdash \Delta} DL$$

$$\frac{\Gamma \vdash \Delta}{\Gamma,\rho^* \vdash \Delta} WL \qquad \frac{\Gamma^* \vdash \rho}{\Gamma^* \vdash \rho} PR$$

In the above rules, Γ, Δ denote a list of regular expressions and ρ,φ denote a single regular expression. In proving $\rho \bullet \varphi$ from Γ, it splits Γ into two disjoint parts.. In dealing with ρ^* construct, the proof system can either discard it, or use ρ at least once.

The above algorithm also tests membership in a regular language: the question of whether a string w is in a regular expression r can be converted to the question of whether w is a subset of r.

3.EXAMPLE

This section describes the use of our algorithm. An example of the use of this construct is provided by the following equivalence: $(l^*)^* = l^*$.This equivalence follows from the facts that $(l^*)^*$ is a subset of l^* and vice versa. These are derived below.

$(l^*)^* \vdash l^*$.

The proof of this as follows:

$l \vdash l$ - Axiom

(1) $l^* \vdash l$ - *DL*
(2) $l^* \vdash l^*$ - *PR*
(3) $l^* \vdash (l^*)^*$ - *PR*

The other direction, $l^* \vdash (l^*)^*$, is proved below.

$l \vdash l$ - Axiom

(1) $l^* \vdash l$ - *DL*
(2) $l^* \vdash l^*$ - *PR*
(3) $l^* \vdash (l^*)^*$ - *PR*

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

As a second example, we will show that $r+s = s+r$,

To do this,we have to show two things.

(a) $(r+s) \supset (s+r)$

(b) $(s+r) \supset (r+s)$

We have a proof of (a) below:

$$\frac{\dfrac{s \vdash s}{r+s \vdash s} + L_2 \dfrac{r+r}{r+s \vdash r} + L_1}{r+s \vdash s+r} + R$$

We have to proof of (b) below

$$\frac{\dfrac{r \vdash r}{s+r \vdash r} + L_2 \dfrac{s+s}{s+r \vdash s} + L_1}{s+r \vdash r+s} + R$$

4.CONCLUSION

We have described an algorithm for testing equivalence of Φ–free regular expressions. The advantage of this algorithm is that it directly captures the real essences regarding language equivalence. As a consequence, it extends easily to other larger language classes such as context-free languages. For example, our algorithm extends easily to the one that deals with algebraic laws, *i.e.*, regular expressions with variables. Two regular expressions with variables are equivalent if whatever expressions we substitute for the variables, the result are equivalent. For example,$\forall L \forall M(L+M = M+L)$. Algebraic laws are a useful tool for simplifying regular expressions. To deal with algebraic laws, it is a simple matter to extend our algorithm with some new rules which introduce universal quantifications...

Regarding the performance of our algorithm, nondeterminism is present in several places of this algorithm. In particular, there is a choice concerning which way the text is split in the•goal. Hodas and Miller[3] dealt with the goal *rs* by using IO-model in which each goal is associated with its input resource and output resource. The idea used here is to delay this choice of splitting as much as possible. This observation leads to a more viable implementation.

Our ultimate interest is in a procedure for carrying out computations of the kind described above. It is hoped that these techniques may lead to better algorithms.

5.ACKNOWLEDGMENT

This paper was supported by Dong- A University Research Fund.

6.REFERENCES

[1] S.C. Kleene, Introduction to Metamathematics, North Holland,Amsterdam,1964.

[2] J.Y. Girard, “Linear logic”, Theoretical Computer Science,vol.50,pp.1-102,1987.

[3] J.Hodas and D. Miller “Logic programming in a fragment of intuitionistic linear logic,” Journal of Information and Computation,1992. Invited to a speical issue of submission to the 1991 LICS conference.

[4] J.E.Hopcroft, R.Motwani, and J.D. Ullman, Automata Theory,Languages and Computation, Addison Wesley, 2006.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

CRS, a Novel Ensemble Construction Methodology

Navid Kardan
Computer Engineering Dep.
IUST
Tehran, Iran
n_kardan@comp.iust.ac.ir

Morteza Analoui
Computer Engineering Dep.
IUST
Tehran, Iran
analoui@iust.ac.ir

***Abstract*— Constructing ensemble classifiers that are both accurate and diverse is an important issue of research and challenging task in machine learning. In this paper, we proposed Class-based Random Subspace (CRS) method; a new ensemble construction method based on the random subspace (RS) strategy, and tested it on a number of standard data sets from UCI machine learning repository. Our results show that CRS is at least as good as RS, and outperforms it in datasets with strong correlation between their classes.**

Random subspace method; Feature selection; Classifier ensemble; Classification

I. INTRODUCTION

The general methods for constructing ensemble classifiers can be applied to any classification algorithm. These methods try to create feature spaces that are as diverse as possible, from the original data set. The most successful of these methods are Boosting [1], Bagging [2], Random subspace [3], and decorate [4]. Another possibly promising approach in this area is the nonlinear transformations of the original data set [5].

In Boosting, every data instance is picked with a probability according to its hardness or misclassification on earlier classifiers of the ensemble. This way the harder instances that need more effort can gain more notification and we can get more diverse ensembles.

In bagging, we make new data sets by sampling with replacement from the original data set. This way we have new data sets that have enough data points and introduce some degree of diversity in them, as well.

In decorate, some artificial instances are added to the original data set for training a new member. The class label of these new instances is set with a value that is different from current ensemble decision. This way we strive for a better community with increasing diversity.

The random subspace method samples features of the original data set, in order to improve diversity, instead of sampling from instances. This way, we have new data sets that have the same amount of data as the original, therefore, can get better classification accuracies. In some applications, such as biometric data processing, this method has proven very good performance [6]. In this paper we improve this approach by taking into account the classes that data items come from. We show that an ensemble created this way, permits to obtain better accuracies than that obtained by a classic RS method.

II. METHODS

An ensemble of classifiers combines different learners in order to obtain a better overall performance. These methods build new data sets from the original data and build the classifiers on these modified data sets. Then the final decision is obtained by combining all of the individual votes.

Suppose a data set with N instances and F features. In the original RS method we build K new data sets, where K is the number of base classifiers in the ensemble, by sampling randomly from the F features and obtain K new training sets each one with N training instances. Then we train each learner with one of these training sets.

The main idea behind this method is the possible redundancy in the original feature space. With sampling from this space we select a subset of the base space and so it is likely to find a space that can lead to better generalization. This improvement in accuracy is obtained from combining these different projections of the feature space. Another advantage of this method is the reduced feature space that can tackle the problem of curse of dimensionality. In fact, Vapnic [7] showed that reducing the feature space is one method for gaining better generalization ability.

Ho in [3] showed that RS can improve accuracy of the tree classifiers considerably, by building a random forest.

RS method treats all features the same. It is also a top-down approach, i.e. it selects the new features at the start of training phase but it does not consider each instance individually. We compensated the original idea of RS according to the fact that the redundancy of data set can be class-based. This means, although there might be better subsets of the feature space that can lead to better classification, but if we consider the data as a whole, we might lose some of these redundancies that are between some parts of the original feature space. These parts of feature space, in our work, correspond to the different classes of the original data set.

For having a better understanding of the situation, consider the case that we have three classes; say A, B, C in the data set. Class A can be best discriminated from other classes with

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

feature subset As and the same is true for classes B and C. in the RS method, we suppose that A_s, B_s and C_s are identical and we can get better results by selecting the subsets that can better approximate the best subset. But in the broader case when these subsets are different, we can get stuck in the situation that leads to no better performance. By taking into account this difference between feature subsets we can get better ensembles and higher accuracies.

Fig. 1 shows a data set with five features and three class types that is sorted according to the class label.

Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	Class label
					Class 1
. . .	. . .	. . .	. . .	. . .	Class 1
					Class 1
					Class 2
. . .	. . .	. . .	. . .	. . .	Class 2
					Class 2
					Class 3
. . .	. . .	. . .	. . .	. . .	Class 3
					Class 3

Fig. 1. A sample data set with five features and three class types.

In figure 2 one possible data set according to RS strategy is presented. In this figure the omitted feature is shown by /// mark. Figure 3 present one possible feature subset according to our method. As depicted in this figure, here we use different feature subsets for each class.

Based on this idea we introduce the idea of Class-based Random subspace (CRS) that can be used to build classifier ensembles that are more likely to have better performance. In CRS every classifier is responsible for one class only and the final decision is obtained by combining different decisions according to their corresponding class.

One important issue about our strategy is the way that we make it possible to have different subsets for different classes. Let we name a specific feature subset as *f*, feature space as *F* and number of class types as *k*. in our method, for each ensemble member we create *k* sub-feature sampling i.e. $m=\{f_1,f_2, \dots , f_k\}$. Each sub-feature f_i will be used to train a classifier that identifies one class. According to this one-vs.-all strategy [8] we will make k classifiers for each unit of our ensemble. If our ensemble consists of l members, $M = \{m_1, m_2, \dots, m_l\}$, we will need $k \times l$ random samples from *F*. In test phase we will count the votes for each class and select the most popular class.

Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	Class label
	///		///		Class 1
. . .	///	. . .	///	. . .	Class 1
	///		///		Class 1
	///		///		Class 2
. . .	///	. . .	///	. . .	Class 2
	///		///		Class 2
	///		///		Class 3
. . .	///	. . .	///	. . .	Class 3
	///		///		Class 3

Fig. 2. Selecting three random features to construct a new data set according to RS strategy

Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	Class label
	///		///		Class 1
. . .	///	. . .	///	. . .	Class 1
	///		///		Class 1
///				///	Class 2
///	. . .	. . .	. . .	///	Class 2
///				///	Class 2
		///	///		Class 3
. . .	. . .	///	///	. . .	Class 3
		///	///		Class 3

Fig. 3. Selecting three random features to construct a new data set according to our strategy

Our algorithm can be described as follows:

- Build *k* data sets from the original data set, where *k* is the number of classes and each set is built for one of these classes; i.e. each set is exactly the original data set except for the class label that can only get two values, one for belonging to the corresponding class and another for belonging to other classes.
- Using each of these new data sets, choose a subset of features randomly and project each instance to this space to build a new data set for each base classifier.
- Train each classifier with its training data and remember its corresponding class.
- For classification of a new instance, any combining method can be used to obtain the final decision of the ensemble. For instance, majority vote can add any vote to its corresponding class and choose the class with most votes.

III. EXPERIMENTS

We used 5 datasets from UCI machine learning repository [9, 10] to evaluate our proposed method. We compared our method with the original RS method for different ensemble sizes and the results show that our method leads to ensembles that are at least as good as the RS and show major improvements in some cases.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

We used Vehicle, Vowel, Soybean, Glass and DNA datasets as our test benchmarks. Vehicle dataset has 846 instances and 19 features including class label. It is used for identifying four types of vehicles from their silhouette. Vowel dataset consists of 990 instances, each with 10 features. It has 11 class types that are the steady state vowels of British English. Soybean dataset has 683 instances with 35 features for each and a class label which may have any of the 19 different values. Glass dataset has 214 observations of the chemical analysis of 7 different kinds of glass. It has 10 features including class label. DNA data set has 181 features including class label and has 3186 instances.

All the experiments are done in R software. Our base classifier is the tree classifier implemented in the tree package of this software [11]. For combining decisions of ensemble members, we used SUM rule [12]. This method computes sum of the decisions of ensemble members for each class and select the class with most value.

For estimating testing error, we used leaving-one-out method [13]. We repeated this procedure for thirty times and averaged the results. Tables I-IV show statistical properties of the number of correctly classified instances for each data set. Each column represents the result of one method for a particular ensemble size. The ensemble size is written in parentheses. Table V represents statistical properties of the two ensemble classification accuracies on the DNA data set.

TABLE I. COMPARISON OF RS AND CRS ON GLASS DATASET USING NUMBER OF INSTANCES THAT CLASSIFIED CORRECTLY

Glass	CRS(5)	RS(5)	CRS(10)	RS(10)	CRS(15)	RS(15)	CRS(20)	RS(20)
Mean	156.9	152.93	159.17	152.13	160.27	151.83	160.43	150.77
MAX	168	163	167	162	167	162	165	156
MIN	150	144	152	146	150	142	154	143
Std. Dev.	4.41	4.62	3.33	4.83	3.62	4.92	2.91	3.57

TABLE II. COMPARISON OF RS AND CRS ON SOYBEAN DATASET USING NUMBER OF INSTANCES THAT CLASSIFIED CORRECTLY

Soybean	CRS(5)	RS(5)	CRS(10)	RS(10)	CRS(15)	RS(15)	CRS(20)	RS(20)
Mean	522.2	513.93	522.4	512.23	525.47	511.87	521.97	511.17
MAX	554	542	534	522	543	521	534	514
MIN	508	506	512	503	511	505	510	504
Std. Dev.	11.49	6.90	6.51	2.88	8.26	3.05	6.54	2.69

TABLE III. COMPARISON OF RS AND CRS ON VEHICLE DATASET USING NUMBER OF INSTANCES THAT CLASSIFIED CORRECTLY

Vehicle	CRS(5)	RS(5)	CRS(10)	RS(10)	CRS(15)	RS(15)	CRS(20)	RS(20)
Mean	614.6	614.47	618.6	616.17	619.87	619.03	623.7	618.53
MAX	638	626	630	626	632	629	636	628
MIN	594	599	606	600	605	600	609	610
Std. Dev.	13.01	7.27	7.57	7.87	6.18	6.85	5.26	4.99

TABLE IV. COMPARISON OF RS AND CRS ON VOWEL DATASET USING NUMBER OF INSTANCES THAT CLASSIFIED CORRECTLY

Vowel	CRS(5)	RS(5)	CRS(10)	RS(10)	CRS(15)	RS(15)	CRS(20)	RS(20)
Mean	788.13	558.23	812.1	574.53	819.37	578.5	825.67	578.8
MAX	809	605	834	662	833	634	843	626
MIN	764	499	788	525	800	524	814	532
Std. Dev.	12.14	30.63	11.54	29.53	9.10	24.58	7.55	21.33

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

TABLE V. COMPARISON OF RS AND CRS ON DNA DATASET USING CLASSIFICATION ACCURACY

DNA	CRS(5)	RS(5)	CRS(10)	RS(10)	CRS(15)	RS(15)	CRS(20)	RS(20)
Mean	0.9341	0.9060	0.9393	0.9086	0.9376	0.9087	0.9386	0.9111
MAX	0.9523	0.9303	0.9517	0.9222	0.9504	0.9253	0.9523	0.9297
MIN	0.9134	0.8901	0.9259	0.8901	0.9165	0.8769	0.9259	0.8883
Std. Dev.	0.0090	0.0101	0.0067	0.0096	0.0094	0.0107	0.0062	0.0096

Figures 4-8 show the average ensemble accuracy vs. ensemble size for each data set. Here we can see that CRS was always at least as good as RS. In Vowel data set, CRS performs much better. In soybean glass and DNA it is considerably better. And in vehicle data set, two methods are not so different.

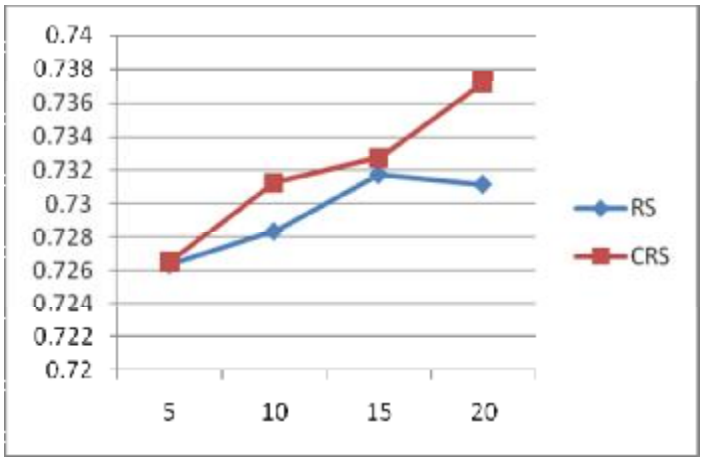


Fig. 4. THE TWO METHODS ARE COMPARED ON VEHICLE DATASET USING ACCURACY VS. ENSEMBLE SIZE

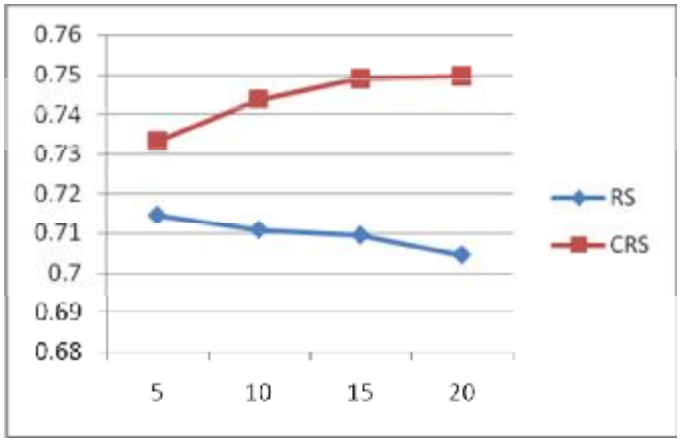


Fig. 5. The two methods are compared on Glass dataset using accuracy vs. ensemble size

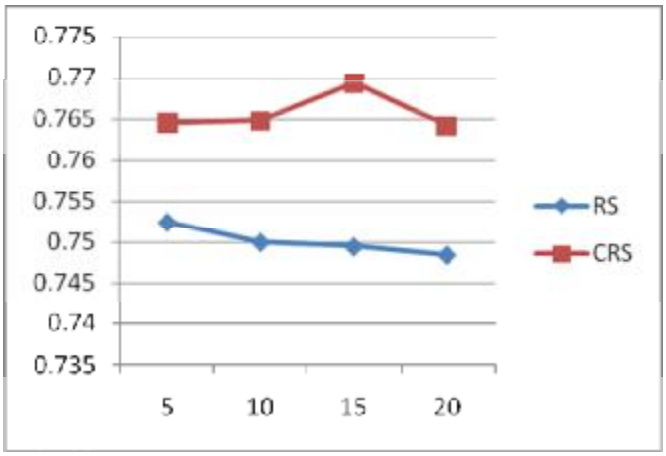


Fig. 6. The two methods are compared on Soybean dataset using accuracy vs. ensemble size

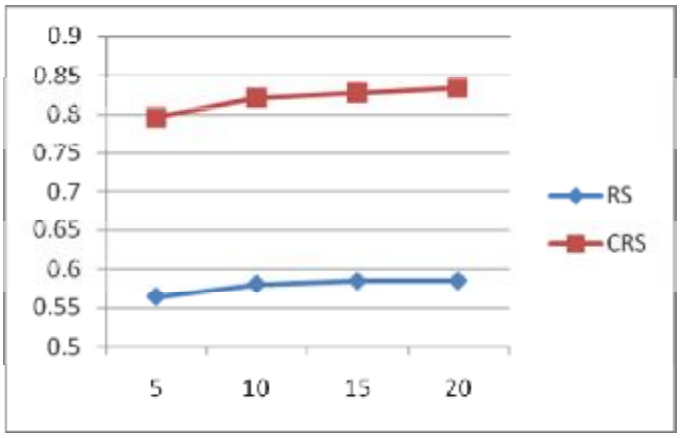


Fig. 7. The two methods are compared on Vowel dataset using accuracy vs. ensemble size

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

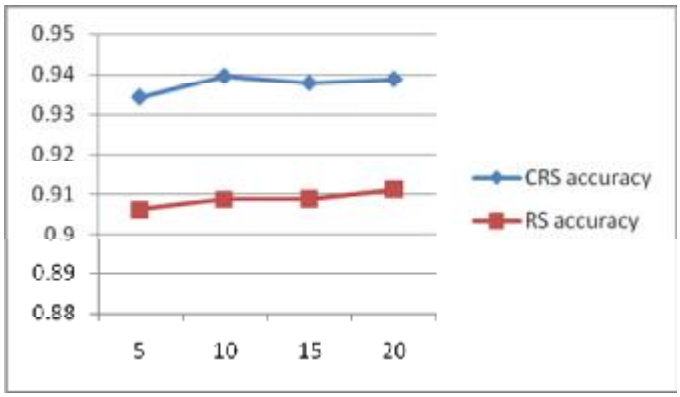


Fig. 8. The two methods are compared on DNA dataset using accuracy vs. ensemble size

IV. CONCLUSION

We proposed a new ensemble construction strategy based on Random subspace method. In this method we tried to compensate classic RS method by trying to reduce the redundancy that might rise from intra-class correlations. If we have a brief comparison with the RS method and the fact that there can be intra-class redundancy in feature space, we can conclude that CRS have the potential to improve overall feature space in the same way that RS can do it in an unsupervised manner. We tested our method on four standard data sets and compared it to the RS method. The results suggest that our method is at least as good as original RS method, but in some cases there is a major superiority. These results are according to the fact that in some data sets there is a stronger correlation between classes and this can lead to better results.

REFERENCES

[1] Y. Freund, R.E. Schapire, "Experiments with a new boosting algorithm", Proceeding of the thirteenth international conference on Machine Learning Bari, Italy, July 3-6, 1996, 148-156.

[2] L. Breiman, "Bagging predictors", machine learning, 24, 1996, 123-140.

[3] T.K. Ho, "the random subspace method for constructing decision forests", IEEE Trans. Pattern Anal. Mach. Intell. 20 (8) (1998) 832-844.

[4] P.Melville, R.J. Mooney, "creating diversity in ensembles using artificial data", in information fusion: Special issue on diversity in multiclassifier systems", vol. 6 (1), 2004, pp. 99-111.

[5] Nicol´as Garc´ıa-Pedrajas, C´esar Garc´ıa-Osorio, Colin Fyfe, "Nonlinear Boosting Projections for Ensemble Construction", Journal of Machine Learning Research 8 (2007) 1-33.

[6] L. Nanni, A. Lumini, "An experimental comparison of ensemble of classifiers for biometric data", Neurocomp. 69 (2006) 1670–1673.

[7] V. Vapnic, "the nature of statistical learning theory", springer-verlag, 1995.

[8] E. Alpaydin, "Introduction to machine learning", Second edition, MIT press, 2010.

[9] D.J. Newman, S. Hettich, C.L. Blake, C.J. Merz, UCI Repository of machine learning databases [http://www.ics.uci.edu/~mlearn/MLRepository.html]. Irvine, CA: University of California, Department of Information and Computer Science (1998).

[10] A. Frank, A. Asuncion, UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science (2010).

[11] B. D. Ripley, "Pattern Recognition and Neural Networks", Cambridge University Press, Cambridge, 1996.

[12] J. kittler, M. Hatef, R. Duin, J. Matas, "On combining classifiers", IEEE Trans. Pattern Anal. Mach. Intell. 20 (3) (1998) 226-239.

[13] R.O. Duda, P.E. Hart, D.G. Stork, "Pattern Classification", second ed., Wiley, New York, 2000.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Routing Optimization Technique Using M/M/1 Queuing Model & Genetic Algorithm

Madiha Sarfraz, M. Younus Javed, Muhammad Almas Anjum, Shaleeza Sohail
Department of Computer Engineering
College of Electrical & Mechanical Engineering
Pakistan
madya.khan@gmail.com, myjaved@ceme.edu.pk, almasanjum@yahoo.com, shaleezas@hotmail.com

Abstract— **Optimization Approaches have been applied to various real life issues in communication and networking. In this research a new approach has been proposed for network path optimization using Genetic Algorithm. The path which is best fitted in the population is considered as the optimal path. It is obtained after qualifying the fitness function measuring criteria. The fitness function measures the best fitted path based on constraints; bandwidth, delay, link utilization and hop count. Population is composition of valid and invalid paths. The length of the chromosome is variable. So the algorithm executes competently in all scenarios. In this paper the comparison of this approach with the fitness function; measuring delay and bandwidth factor, has also been catered. This work has been performed on smaller network; work is in progress on large network. Thus, the results proved our affirmation that proposed approach finds optimal path more proficiently than existing approaches.**

I. INTRODUCTION

OPTIMIZATION in the field of Genetic algorithm is gaining massive magnitude. GA is globally used optimization technique [1] based on natural selection phenomenon [2]. It is considered as an important aspect in networking. From source node to destination node essential solution is optimization which is needed. The network traffic flow is mounting rapidly. So the balance in Quality of Service and Broadcast of traffic should be maintained. Increased in load of network traffic will cause the delay in traffic and affect the QoS as well. The routing problem scenarios can be resolved through optimization [3].

In the research the approach which has been introduced is of network optimization routing strategy using Genetic Algorithm. It involves bandwidth, delay and utilization constraints. All these will be in different catering in scenarios. It digs out the most optimal path from the population lot on the basis of Fitness Function. The fitness function selects the path which has less delay, less utilization factor, more bandwidth availability and less number of hops to be travel. The hop count is used as the decision making factor when there are more than one path with same strength. The chromosome represents the path and the population is showing collection of feasible and infeasible paths. Each chromosome has variable number of nodes. This routing strategy is not efficient and robust merely but it also congregates swiftly.

The paper is formatted as: The related work is done under section II. Overview of Genetic Algorithm has been structured in section III. In this the outline of GA and its operators are explained. Section IV explains the optimization strategy for routing using GA. In this two fitness functions are compared. The proposed fitness function includes bandwidth, delay, utilization and hop count. It is compared with the fitness function which is catering bandwidth and delay factor. The results and analysis is illustrated in section V.

II. RELATED WORK

Genetic Algorithm which is a versatile technique designed for optimization and searching network planning control in Anton Riedl [4] work. It is also in planning of integration of packet switched network.

Yinzhen Li, Ruichun He, Yaohuang Guo [20] work on finding out optimal path with fixed length chromosomes. This is following priority-based mechanism. In this work Mitsuo. Gen technique's loophole has also been indicated.

In Carlos A. Coello Coello's tutorial [26] multiple objectives optimization has been discussed. The pros and cons of these approaches are done by them. Also illustrate research done and their implications in the respective field.

Introduction of key-based technology in optimization is done by Mitsuo Gen and Lin Lin [16]. They have combined different operators [16]. They high level of search paradigm leading to improvement in computational time and path optimization.

Basela S. Hasan, Mohammad A. Khamees and Ashraf S. Hasan Mahmoud [24] used Heuristic mechanism for Genetic Algorithm. They have considered single source shortest path. For searching heuristic approach is used for crossing over also

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

called as recombination and mutation.

In other work of Anton Riedl [19] title represents the research done is upon path optimization in traffic engineering scenario. It discusses the implications of network optimization and traffic engineering. It main focus is routing with multiple delay and bandwidth constraints for optimization.

Andersson and Wallace [25] proposed a GA which is robust and requires few numbers of parameters. It main emphasis is how multiple objectives GA works out on real life scenarios.

An approach has been developed for reducing congestion in the network [17] by M. Ericsson, M.G.C. Resende and P.M.Pardalos.

Ramon Fabregat, Yezid Donoso, Benjamin Baran, Fernando Solano and Jose L. Marzo [23] presented traffic-engineering load balancing classification. They have not work on packet loss and other factors like backup paths. They introduced GMM model.

Diverse and versatile genetic algorithm is proposed in the work of Norio Shimamoto, Atsushi Hiramatsu and Kimiyoshi Yamasaki [18]. In this approach last result of the iteration is used for the next generation. Performance level of the algorithm is good.

Abdullah Konak, David W. Coit, Alice E. Smith [21] shows that investigation for solutions is the coherent response for multiple objectives. A real life scenario entails immediately multiple objectives for optimization. It also give overview of GA which are developed for multipurpose objectives and maintaining the diversity. They have introduced GR (Greedy Reduction) technique. In the case of worst scenario GR technique executes in the linear time of framework [22].

For bandwidth allocation Hong Pan and I. Y. Wang [15] proposed GA for optimization. In this the average delay network is lessened. Bandwidth can be reassigned again as per this algorithm.

III. Overview Of Genetic Algorithm (GA)

Genetic Algorithm is a search paradigm. It follows principles which are based on Darwin Theory of evolution. In this population data fights for survival and the 'fittest' one survives. This algorithm is mainly based on natural selection phenomenon. GA is introduced by John Holland [7] [5]. It is effective technique as its not only encountering mutation but also use crossing over technique (or genetic recombination) [10]. Crossing over technique improves the proficiency of algorithm for having the optimum outcome. Holland had a dual aim [11]:

- To improve the concept of understanding of natural selection process
- To develop artificial structure having functionality analogous to natural system

Genetic Algorithm has been broadly used for solving MOPs because it works on a population of solutions. [13]. Objective function plays an important role for obtaining optimum result. For having "Optimum" result does not mean that the result is "Maximum". It's the best and the most appropriate value as per the objective function criteria. Genetic algorithm is best for optimization and is useful in any state.

A. Population

The Genetic algorithm starts up with a set of solutions which is taken as a 'population' with the assumption that next generation will be better than the previous one. The cycle is terminated when the termination condition is satisfied.

B. Operators of GA

The efficiency of Genetic algorithm is dependent on the way the operators are used, that constitutes Genetic algorithm course of action [9]. The GA operators are as follows [8] [4]:

1) Selection & Reproduction: Chromosomes are selected according to their Objective Function (also called as fitness function). The ultimate node for chromosome survival is Objective Function. It works on Darwinian Theory for survival of the fittest. This is an artificial version of natural selection for survival. Chromosomes having higher fitness have greater likelihood of being into next generation. There are numerous methods for chromosome selection. For selection of chromosomes several methods are in practice.

2) Crossover or Recombination: It is the distinguishing feature from other techniques of optimization. On selected parents the technique is applied as per their crossing-over mechanism.

- 1-point Crossing over
- 2-point Crossing over

Crossing over results into new off-springs.

3) Mutation: In GA maneuvering mutation has secondary role [6]. It is required after crossing over segment because there is a probability of information loss at this stage [12]. It is done by flipping by bit as per requirement. The population diversity maintenance is purpose of mutation operator.

C. Steps to follow in GA

1) Population Generation: Generate population randomly of 'n'.

2) Fitness Function: Launch function for evaluation of fitness in population.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

3) Applying Operators: Create new population by applying operators of GA until the new population is complete. At the end of iteration another generation is attained. The Rank Based Selection is applied.

a) Selection: Parents selection is on the basis of their fitness function creteria.

b) Recombination: Recombination also called as crossing over is than applied over the selected parents. It result into offsprings.

c) Mutation: The resulting offsprings are than mutated. After that their fitness function is measured. On the basis of this value, their survival in the population is based.

4) Terminating Condition: If the terminating condition is reached, the loop breaks and the best result is obtained.

5) Resulting Generation: Newly generated population is used for further generation.

IV. Proposed Optimization Routing Approach

In this proposed optimization approach basic unit of a chromosome is gene. Genes constitute to form a chromosome, which in turn constitutes the population. Thus the chromosome in population is represented by string of number as in Figure 1.

Figure 1. Genetic Representation of Path

The gene represents the node while the chromosome represents the network path. Population is the collection of all possible paths. The chromosome 1-5-3-6 shows network path which is constituted by nodes. The first node is the source node while the last node is the destination node. The hop count in this will be:

Hop Count = Chromosome Length - 1

The length of the chromosome is variable. So the number of nodes in the path is directly proportional to number of hop the data has to travel. The chromosome which have source and destination same as defined for the path, they constitute the population lot. The rest unfeasible paths will be discarded after they are generated. From those feasible paths, the population is generated randomly.

A. Strategies

The proposed strategy is explained in Fitness Function II and is compared with Fitness Function I as well. The proposed strategy is more efficient as it is handling more constraints than the other approach [27].

B. Fitness Function I

This fitness function is for finding the delay of the path using the delay and bandwidth constraints [27]. The algorithm is as follows:

- It checks the bandwidth availability of link.
- After that network delay of link is find out
- Than delay average of path is calculated.

All the paths are valid, as they have passed through the bandwidth availability check before calculating the average packet delay as [27]:

$$AveragePacketDelay = \frac{delay}{no.\ of\ links} \quad (1)$$

The delay of the path is calculated as [27]:

$$delay = \sum_{i=1}^{n} \frac{bwAvailable_i}{DataSize} \quad (2)$$

The optimum path selection is based mainly on the bandwidth and delay in this approach. Thus, if two or more paths have same fitness value than the path with less number of hops will be taken as the optimum path. Optimum path leads to a path which has more bandwidth, less delay and less number of hops.

Now the proposed strategy involving is illustrated as follows:

C. Fitness Function II

The proposed strategy is founded on M/M/1 queuing model [14] by using GA for Path Optimization. In our research it is used for handling Bandwidth, Utilization and Delay constraints for finding Optimum path. The fitness function is based mainly on (3). If there is a conflict in making decision among these constraints than path with lesser hops is taken as the optimum path. Constraints defined are Bandwidth Availability, Delay, Utilization and Hop count. The fitness function using M/M/1 queuing model [14] is as follows:

$$T = \frac{1/\mu C}{1-\rho} \quad (3)$$

T = Delay mean (sec)
1/μ = mean packet size (bits)
C = capacity (bps)
ρ = utilization factor

1) Capacity 'C': It is the bandwidth C available over the network path.

2) Utilization factor 'ρ': After checking of bandwidth availability, the network utilization factor is obtained as:

$$utili_i = \frac{DataSize}{bwAvailable_i} \quad (4)$$

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

After summing up of utilization factor a links, utilization factor value for a single path is obtained. Than mean of the utilization factor of path is obtained by dividing the sum of the utilization factor value for a single path with the total number of links in that path as:

$$Mean\ Utilization = (\sum_{i=1}^{n} Utilization)/total\ number\ of\ links \quad (5)$$

3) Delay Mean: The equation (3) is precisely formulized as follows:

$$\text{Delay Utilization Mean} = (Mean\ Utilization)/ (1- Mean\ Utilization) \quad (6)$$

It constitutes fitness function of this algorithm:

$$\text{Fitness Function} \sim \text{Delay Utilization Mean} \quad (7)$$

This utilization factor is calculated for every path. After each generation the mean of all paths is also calculated. All the paths are valid path as their bandwidth availability is checked before finding out their fitness. Optimum path selection is based on the link utilization, Bandwidth, Delay Factor and hop count. The more the bandwidth, the less the delay, utilization and the hop count the more optimum the path will be. When there are two or more paths with same fitness value than hop count will be playing the decision making role.

D. Proposed GA for Path Optimization

Proposed Routing Strategy using GA for network path optimization is as follows:

*1) Initialization of Population:*The population is randomly generated [27]. There also exists the probability of valid and invalid paths. The source and destination nodes are fixed. The randomly selected population is 33% of the generated chromosomes. So this 33% constitutes the whole population and 30 generations are produced [27].

2) Selection of Parents: Parent's selection is based on Rank-based Selection. In this the parents are ranked on the basis of their fitness function value. They are not selected randomly except the first generation (in which parents are selected randomly). The parents with high fitness will be higher in ranking and the ones with low fitness will be lower in ranking. Both parents which are best in the population are used for crossing over as they have more tendency of survival in nature.

3) Recombination (or Crossing Over): Crossing over of the best fitted chromosome results into best offsprings. The crossing over technique used is "2-point over 1-point crossover" [27]. In this first 2-point crossing over technique is applied on the selected chromosomes, leading towards the segregation of source and destination node. After this 1-point crossing over technique is done over the resultant chromosome. In this the crossing over point is 2 [27].

4) Mutation: After crossing over, mutation is being done on the offsprings. Mutation has been done as per the scenarios which are [27]:

a) Scenario 1 (repeating node):

- The location of repeating node is traced out.
- Any of the missing nodes is find out.
- Place the missing node at traced location.

b) Scenario 2 (no missing or repeating node):

- Randomly pick two nodes from chromosome
- Swap those nodes

c) Scenario 3 (minimum chromosome length):

- Minimum length of chromosome is 2
- It has only source and destination nodes
- There will be no mutation

d) Scenario 4 (length of chromosome is one more than the minimum):

- Length of chromosome one more than minimum constitutes of 3 nodes
- Only middle node is flipped with missing node

5) Evaluation of Mutated Offsprings: The mutated offspring's fitness is evaluated. The chromosome with worst fitness is replaced by the offspring having better fitness. The population size will remain unaffected by this replacement. The worst chromosomes are discarded. The network path survival is based on their Fitness Function Criteria.

V. Experiments, Results & Analysis

A. Experiments and Testbed

1) Network Formation: The delay, bandwidth and utilization factor has been evaluated on the following network:

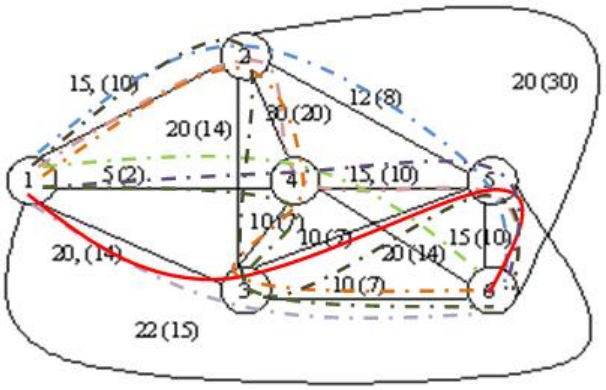


Figure 2. Network Formation

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Main factors of algorithm are availability of bandwidth, utilization, delay and hop count in network traffic. When the packet is passed through a network, two parameters are involved. One is time packet takes to reach next node and other is bandwidth. Time also shows distance packet travels. The variety of time and bandwidth measures is included in the network, leading to better result in a complicated network.

2) Population Selection Criteria: The population is selected randomly. The result of randomly selected population's result is compared with whole population and with results of the [15] also. The obtained results are better than both of them. Hence, proved that random selection is best choice [27]. The criterion for population size and population generation is as shown in the following table [27]:

TABLE I. TABLE SHOWING POPULATION SIZE & NUMBER OF GENERATIONS

Population Size	33%
Generations	30

3) Experimental Practice: The population is selected randomly. First bandwidth availability is checked and then calculates utilization and delay factor as mentioned in fitness function. The hop count is being considered for resolving the conflict between paths with same fitness value. The paths with less hop count in that case will be considered as an optimum path. Even if the hop count is same, than select any of the path which is not used often.

B. Analysis & Results

Fitness I is applied over the whole population. The population is generated 50 times. The minimum average Delay is taken till 50 iterations and is shown in the graph below. So, there will be a delay in getting the optimum result.
The horizontal axis is showing the generation and the vertical axis is showing the delay average (fitness function value) against every generation.

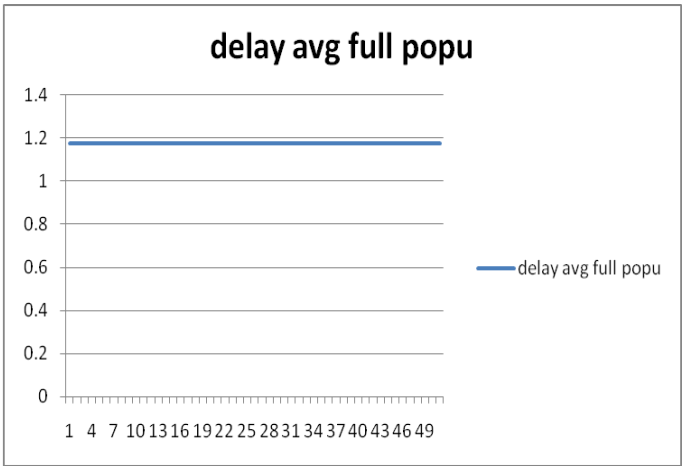


Figure 3. Average Delay (whole population)

In this case number of hops is constant. There is no alteration in hop count till the 50 generation. It might show change later. The graph is constant showing no variation.

Now the population is randomly selected and fitness function I is implemented over it. The results are compared with the full population selection results. The 33% of the population is taken and is generated 30 times as per Table I. The minimum of delay average of every generation is shown in the Figure 4. Hence, it is showing better results than shown in the Figure 3. There is a variation in this graph and is reaching the minimum level of delay.

The horizontal axis is showing the generation and the vertical axis is showing the delay average (fitness function value) against every generation with random selection of population.

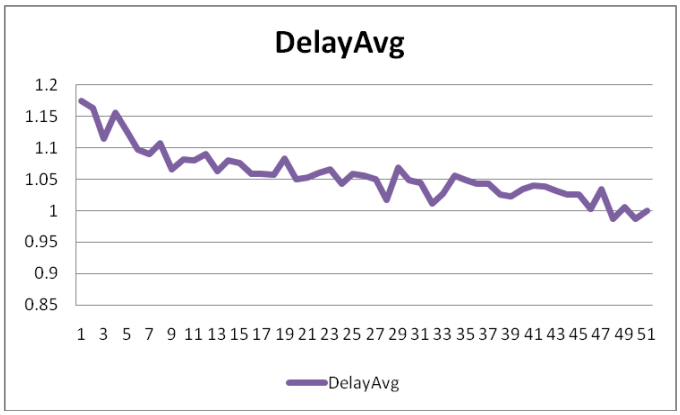


Figure 4. Average Delay (randomly selected population)

As per the results obtained the number of hops is also reducing. The hops are more at the start of the generation, but as the generations are increasing their number of hops is also reducing. Therefore, it is involving traffic engineering of the network as well. So in this case it is sharing the load of network through delay and bandwidth constraints. In this utilization factor is not involved in fitness function criteria.

At start the hops count of the paths are more as per the network scenario, but in final result the number of hops is lesser. Thus, it is also handling the traffic engineering problem of the network. This is thus sharing and balancing the network load as per bandwidth and delay criteria.

Now the proposed strategy (fitness function II) is applied over the network. Population selection is random [27]. It involves the delay, bandwidth and utilization constraints. It also involves hop count.

The horizontal axis is showing the generation and the vertical axis is showing the delay and utilization measure (fitness

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

function II) against every generation with random selection of population.

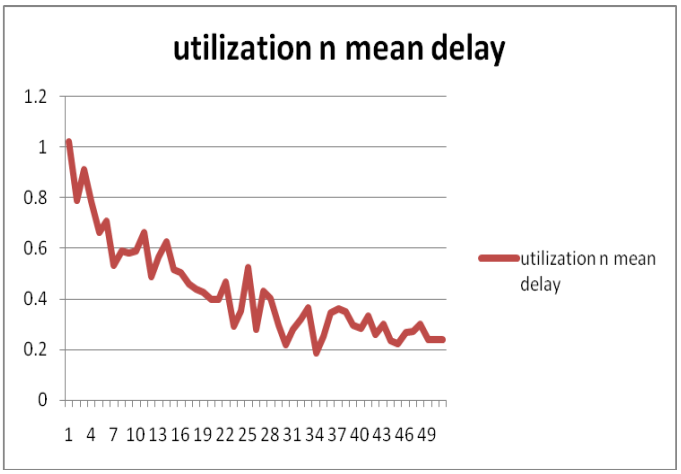


Figure 5. Proposed Algorithm involving utlization and Delay constraints with randomly selected population

In this graph there is varitaion in fitness value involving delay and utilization factors of the network. Thus it is reaching the minimum value more earlier than the previous algorthims discussed. It is reaching getting value below 1 at ealier stage which fitness function I graph is showing after mid generation.

In Figure 6 the fitness function which is involving more constraints are showing better results than compared to the other fitness function results mentioned in this research paper. So for calculating the optimum path of the network the Fitness Function II should be applied over the network for congestion control and traffic engineering problems.

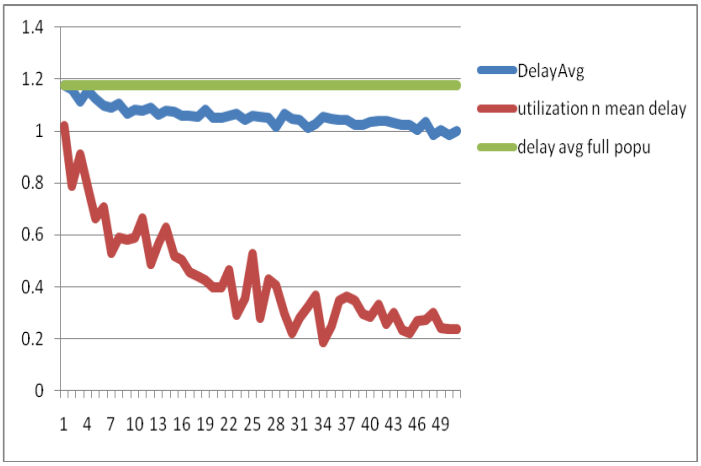


Figure 6. Comparison of both fitness function and respective population selected (whole and random selection)

Figure 6 is showing comparison of the proposed algorithm with the other algorithm mentioned in this research.

VI. Conclusion

Bandwidth scaling is elementary driver of popularity and growth among interconnected computer networks. Increase in bandwidth accompanies lesser delay. In this research, the proposed technique is for finding optimized path with delay, bandwidth and utilization measures. In this the comparison is done between the fitness function catering just bandwidth and delay with the fitness function handling bandwidth, delay and utilization. Both of them are tackling the hop count. The results prove our affirmation that proposed algorithm shows better results than the earlier proposed algorithms.

References

[1] David E. Goldberg, "Genetic Algorithms in Search, Optimization & Machine Learning", Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA, 1989.

[2] A.W.W.NG and B.J.C. Perera, "Importance of Genetic Alrothim Operators in River Water Quality Model Parameter Optimisation", school of the Built Environment, Victoris University of Technology, Melbourne.

[3] Zhao-Xia Wang, Zeng-Qiang Chen and Zhu-Zhi Yuan, "QoS routing optimization strategy using genetic algorithm in optical fiber communication networks", Journal of Computer Science and Technology, Volume 19 , Year of Publication: 2004, ISSN:1000-9000, Pages: 213 - 217

[4] Anton Riedl, "A Versatile Genetic Algorithm for Network Planning", Institute of Communication Networks Munich University of Technology.

[5] Darrell Whitley, "A genetic algorithm tutorial", Statistics and Computing, 1992, volume 4, number 2, pages 65-85.

[6] David E. Goldberg, "Genetic Algorithms in Search, Optimization & Machine Learning", Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA, 1989.

[7] Darrell Whitley, "A genetic algorithm tutorial", Statistics and Computing, 1992, volume 4, number 2, pages 65-85.

[8] David E. Goldberg, "Genetic Algorithms in Search, Optimization & Machine Learning", Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA, 1989.

[9] A.W.W.NG and B.J.C. Perera, "Importance of Genetic Alrothim Operators in River Water Quality Model Parameter Optimisation", school of the Built Environment, Victoris University of Technology, Melbourne.

[10] Emmeche C., "Garden in the Machine. The Emerging Science of Artificial Life", Princeton University Press, 1994, pp. 114.

[11] Goldberg D., "Genetic Algorithms", Addison Wesley, 1988.

[12] David M. Tate, Alice E. Smith, "Expected Allele Coverage and the Role of Mutation in Genetic Algorithms", Proceedings of the 5th International Conference on Genetic Algorithms Pages: 31 - 37 Year of Publication: 1993, ISBN:1-55860-299-2

[13] Dinesh Kumar, Y. S. Brar, and V. K. Banga, Multicast Optimization Techniques using Best Effort Genetic Algorithms, World Academy of Science, Engineering and Technology 50 2009

[14] Yantai Shu, Fei Xue, Zhigang Jin and Oliver Yang, " The Impact of Self-similar Traffic on Network Delay", Dept. of Computer Science, Tianjin University, Tianjin 300072, P.R. China, School of Information Technology and Engineering, University of Ottawa,Ottawa, Ontario Canada KIN 6N5, 0-7803-43 14-X/98/$10.0081998 IEEE

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

[15] Hong Pan, Irving Y. Wang, "The Bandwidth Allocation of ATM through GA", Global Telecommunications Conference, 1991. GLOBECOM '91, 1991, pages 125-129.

[16] Mitsuo Gen, Lin Lin , "A New Approach for Shortest Path Routing Problem by Random Key-based GA", Genetic And Evolutionary Computation Conference Proceedings of the 8th annual conference on Genetic and evolutionary computation, 2006.

[17] M. Ericsson, M. Resende, P. Pardalos, "A Genetic Algorithm for the weight setting problem in OSPF routing", Journal of Combinatorial Optimization, 2002, volume 6, pages 299-333.

[18] Norio. Shimamoto, Atsushi Hiramatsu, Kimiyoshi Yamasaki, "A dynamic Routing Control based on a Genetic Algorithm", IEEE International Conference on Neural Networks, 1993, pages 1123-1128.

[19] Anton Riedl, "A Hybrid Genetic Algorithm for Routing Optimization in IP Networks Utilizing Bandwidth and Delay Metrics", Institute of Communication Networks, Munich University of Technology, Munich, Germany.

[20] Yinzhen Li, Ruichun He, Yaohuang Guo, "Faster Genetic Algorithm for Network Paths", Proceedings of The Sixth International Symposium on Operations Research & Its Applications (ISORA'06) Xinjiang, China, August 8-12, 2006. Pages 380-389.

[21] Abdullah Konak, David W. Coit, Alice E. Smith, "Multi-objective optimization using genetic algorithms: A tutorial"

[22] Vandana Venkat, Sheldon H. Jacobson and James A. Stori, "A Post-Optimality Analysis Algorithm for Multi-Objective Optimization", Journal Computational Optimization and Applications, Volume 28, 2004.

[23] Ramon Fabregat, Yezid Donoso, Benjamin Baran, Fernando Solano, Jose L. Marzo, "Multi-objective optimization scheme for multicast flows: a survey, a model and a MOEA solution", Proceedings of the 3rd international IFIP/ACM Latin American conference on Networking, 2006, pages 73-86.

[24] Basela S. Hasan, Mohammad A. Khamees, Ashraf S. Hasan Mahmoud, "A Heuristic Genetic Algorithm for the Single Source Shortest Path Problem", IEEE/ACS International Conference on Computer Systems and Applications, 2007, pages 187-194.

[25] Johan Andersson, "Applications of a Multi-Objective Genetic Algorithm to Engineering Design Problems", Department of Mechanical Engineering, Linköping University 581 83 Linköping, Sweden

[26] Carlos A. Coello Coello, "A Short Tutorial on Evolutionary Multiobjective Optimization"

[27] Madiha Sarfraz, Younus Javed, Almas Anjum, Shaleeza Sohail, "Routing Optimization Strategy Using Genetic Algorithm Utilizing Bandwidth and Delay Metrics", 2010 the 2nd International Conference on Computer and Automation Engineering (ICCAE 2010), Singapore

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Architectural Description of an Automated System for Uncertainty Issues Management in Information Security

Haider Abbas
Department of Electronic Systems,
Royal Institute of Technology, Sweden
haidera@kth.se

Christer Magnusson
Department of Computer and System Sciences,
Stockholm University, Sweden
cmagnus@dsv.su.se

Louise Yngström
Department of Computer and System Sciences,
Stockholm University, Sweden
louise@dsv.su.se

Ahmed Hemani
Department of Electronic Systems,
Royal Institute of Technology, Sweden
hemani@kth.se

Abstract*— Information technology evolves at a faster pace giving organizations a limited scope to comprehend and effectively react to steady flux nature of its progress. Consequently the rapid technological progression raises various concerns for the IT system of an organization i.e. existing hardware/software obsoleteness, uncertain system behavior, interoperability of various components/methods, sudden changes in IT security requirements and expiration of security evaluations. These issues are continuous and critical in their nature that create uncertainty in IT infrastructure and threaten the IT security measures of an organization. In this research, Options theory is devised to address uncertainty issues in IT security management and the concepts have been deployed/validated through real cases on SHS (Spridnings-och-Hämtningssystem) and ESAM (E-Society) systems. AUMSIS (Automated Uncertainty Management System in Information Security) is the ultimate objective of this research which provides an automated system for uncertainty management in information security. The paper presents the architectural description of AUMSIS, its various components, information flow, storage and information processing details using options valuation technique. It also presents heterogeneous information retrieval problems and their solution. The architecture is validated with examples from SHS system.

Keywords: *Information Security, Uncertainty Issues, Options Theory*

I. INTRODUCTION

Technological uncertainty due to rapid development and innovation in IT, continuously impacts security measures of an organization. The development is desirable that could facilitate business organizations with innovative hardware, novel methods and state of the art technologies. While on the other hand, technological progression also requires business organizations to adapt new methods and technologies to secure their information system (storage, retrieval, communication etc) processes. The objective could be achieved by deploying new security methods and by evaluating their validity, serviceability and interoperability using re-evaluation. But the service acquisition and validation process for IT security mechanisms is victimized by uncertainty due to new unforeseen threats and technological advancements appearing from time to time. Also these newly acquired security services/features may affect other interacting systems, this is referred to as externalities [1][2]. We addressed three major concerns in information security management due to technological uncertainty i.e. dynamically changing security requirements [3], IT security externalities [4] and obsoleteness of security evaluations [5]. We have utilized options theory from corporate finance [6]; known due to significance of providing effective guidance during uncertain investments. The options theory has been transformed using adaptability model [7] to tailor the IT security processes. The options theory methods were manually applied to illustrate and validate the concepts using real cases on ESAM (E-Society) [8] and SHS (Spridnings-och- Hämtningssystem) [9] systems. The ultimate objective of this research is to develop an automated solution (AUMSIS: Automated Uncertainty Management System in IT Security) for uncertainty issues management in IT security. The solution can be deployed in an organization and will be capable of providing system generated reports for; i) requirement change summary and suggested solutions ii) externalities report and internalization parameters and iii) re-evaluation strategy/guidance based on actual system state. In this paper, we will present the architectural description of the AUMSIS system which consists of its various components, architectural styles, information flow between components, storage details and heterogeneous information processing description.

The paper is organized as follows: Next in section 2, the related work will be highlighted, section 3 presents the holistic view of the IT security uncertainty issues and section 4

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

presents the concept of automated uncertainty management solutions and elaborates its various constituents. Section 5 describes the information processing and flow in AUMSIS. Section 6 elaborates heterogeneous information processing problem and the proposed solution for this issue. Section 7 presents the discussion about the analysis and validation of the AUMSIS framework using SHS example. Section 8 presents conclusion and the future intention of this research.

II. Related Work

Automated information processing systems have been emphasized from various researchers in many domain areas. For example, Wilson, D. et al. has discussed various issues in automated inspection and representation of uncertainty for the real world issues [10]. McVicker, M. et al. has presented the infrastructure that collects statements of security-related statistics from the World Wide Web for source reliability verifications [11]. The work presented in this paper addresses the automated solution of uncertainty issues that might suddenly appear during IT security requirements/evaluation management and require a cumbersome solution exploration process with significant resources [12]. The ultimate outcome of this research will benefit organizations to have system-generated reports for IT security management i.e. i) changing requirements solutions ii) internalization guidance, iii) re-evaluation strategies and iv) security investment related suggestions/decisions.

III. Information Security and Uncertainty Issues

Most of the businesses today rely on IT infrastructures and have to deploy various security mechanisms to protect their work processes. Technological uncertainty strongly impacts those security mechanisms, which become obsolete with the rapid technological progression. The research emphasizes three critical concerns caused by technological uncertainty for an organization in IT security perspective as depicted in Figure 1.

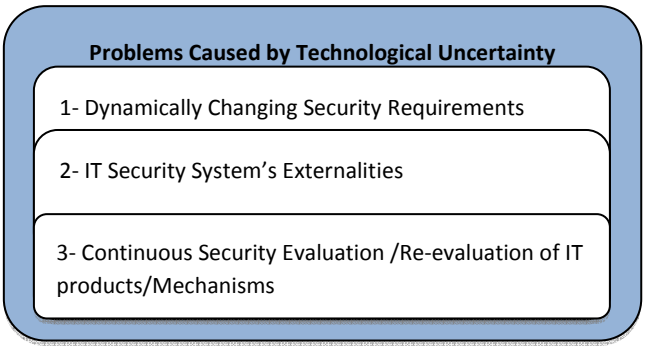


Figure 1. Uncertainty issues addressed in AUMSIS

An organization continuously has to go through a cumbersome procedure to deal with uncertainty issues and to keep their IT system up-to-date and according to new technological standards. The research aims for an infrastructure that will help to avoid the resource-hungry procedures and frame the system state, organizational needs, system's externalities issues and re-evaluation requirements analysis. The next section presents the architectural details of such an automated system (AUMSIS) that can be deployed in an organization. The system will automatically generate uncertainty solution reports for the issues depicted in Figure 1.

IV. Automated Uncertainty Management Solutions in Information Security

AUMSIS is aimed to provide system-generated strategic guidance for above-mentioned issues described in section III. Decision-makers can use this information to formalize current and future IT security management strategies based on actual system state, which consists of organizational policies, up-coming technologies, vulnerability logs, attack history and available budget. Figure 2 depicts the abstract view of the AUMSIS architecture as follows:

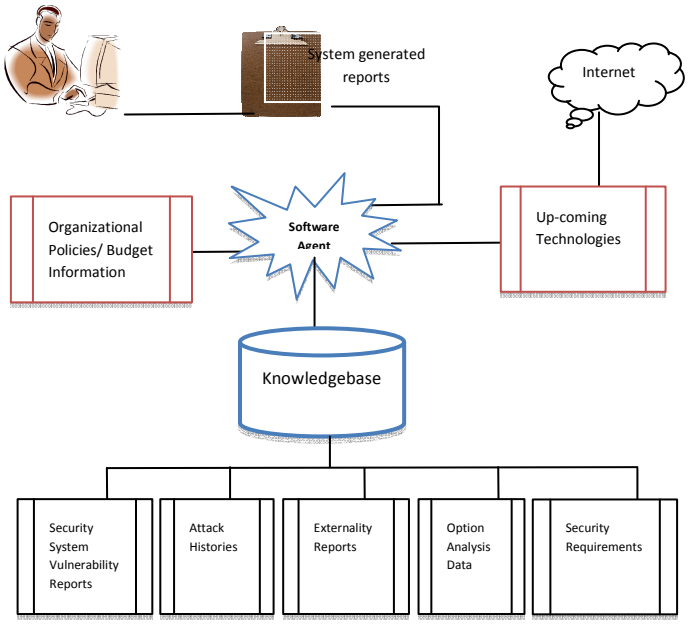


Figure 2. AUMSIS Architecture

The various components of AUMSIS architecture depicted in Figure 2 are elaborated as follows:

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

A. Knowledgebase

Information related to system state during a specified time period is named as historical data and organized in a structured repository; knowledgebase. It consists of following components:

i) System vulnerability reports

It contains malfunctioning reports of the security system and the corresponding affected security components. The information can be used to track the actual service/component causing vulnerability and provides details to determine system state.

ii) Attack history

Attack history data contains information about the exploitation of a particular security service/component by authorized/unauthorized sources. It will reveal shortcomings in existing security mechanisms that need to be factored in.

iii) Externality reports

IT security system of an organization may also leave positive or negative effects to other interacting systems/sub-systems referred to as externalities [1]. Externalities of a security system [2] can be identified by internal/external malfunctioning reports from affected systems/partners. Externality reports provide a holistic view of the IT security system and help to determine system's desired functionality.

iv) Options analysis data

AUMSIS generates results using options technique that are reusable by subsequent analysis. Options analysis data contains information about already executed options and results from a previous analysis. Option cards were used to store data about the options analysis outcomes [3].

v) Security Requirements

Security requirement change reports from various stakeholders or security requirements from any external enforcing authority. This is continuously updated to factor in new/changed requirements.

B. Up-coming Technology and Organizational policies/Budget information

It is the prime objective of AUMSIS to provide contemporary guidance about requirement solutions, internalization factors and evaluation strategy. Therefore the AUMSIS has to interact and extract information from Internet and organization's policy database. These factors are considered as a separate component in AUMSIS due to their evolving nature during analysis.

Above-mentioned historical data, information about upcoming-technologies and organizational policies/budget are accumulated over time and readily available for options analyst agent (OAA) for processing.

C. Options Analyst Agent (OAA)

Options analysis agent is a piece of software [13] that formalizes requirement solutions, internalization results and evaluation strategy using options technique. It extracts system state from knowledgebase, Internet (for up-coming technologies) and organizational policy/budget database. OAA generates the strategic information for decision makers i.e.

i) Analyze alternative solutions for a security requirement and provides recommendations based on contemporary system state.

ii) Internalizing solutions for externalities according to organizational policies.

iii) Deterministic test plans strategies for the evaluation process of each security service considering its malfunctioning report and service exploitation history.

AUMSIS provides up-to-date strategic guidance for the uncertainty issues in information security management process. It considers uncertainty elements caused by changing environment and helps to devise respective optimal IT security strategy. Next section describes the information processing and flow in AUMSIS system.

V. AUMSIS INFORMATION PROCESSING AND FLOW

AUMSIS provides strategic guidance for three main areas of information security management affected by uncertainty issues. The uncertainty management process using AUSMIS for these issues follows slightly different mechanism due to the nature of problems. But the data are maintained in a single repository i.e. knowledgebase. As the AUMSIS addresses three uncertainty concerns in IT security, each one is elaborated individually in Module 1, Module 2 and Module 3. Figure 3 below depicts the information flow of these three modules using information flow diagram as follows:

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

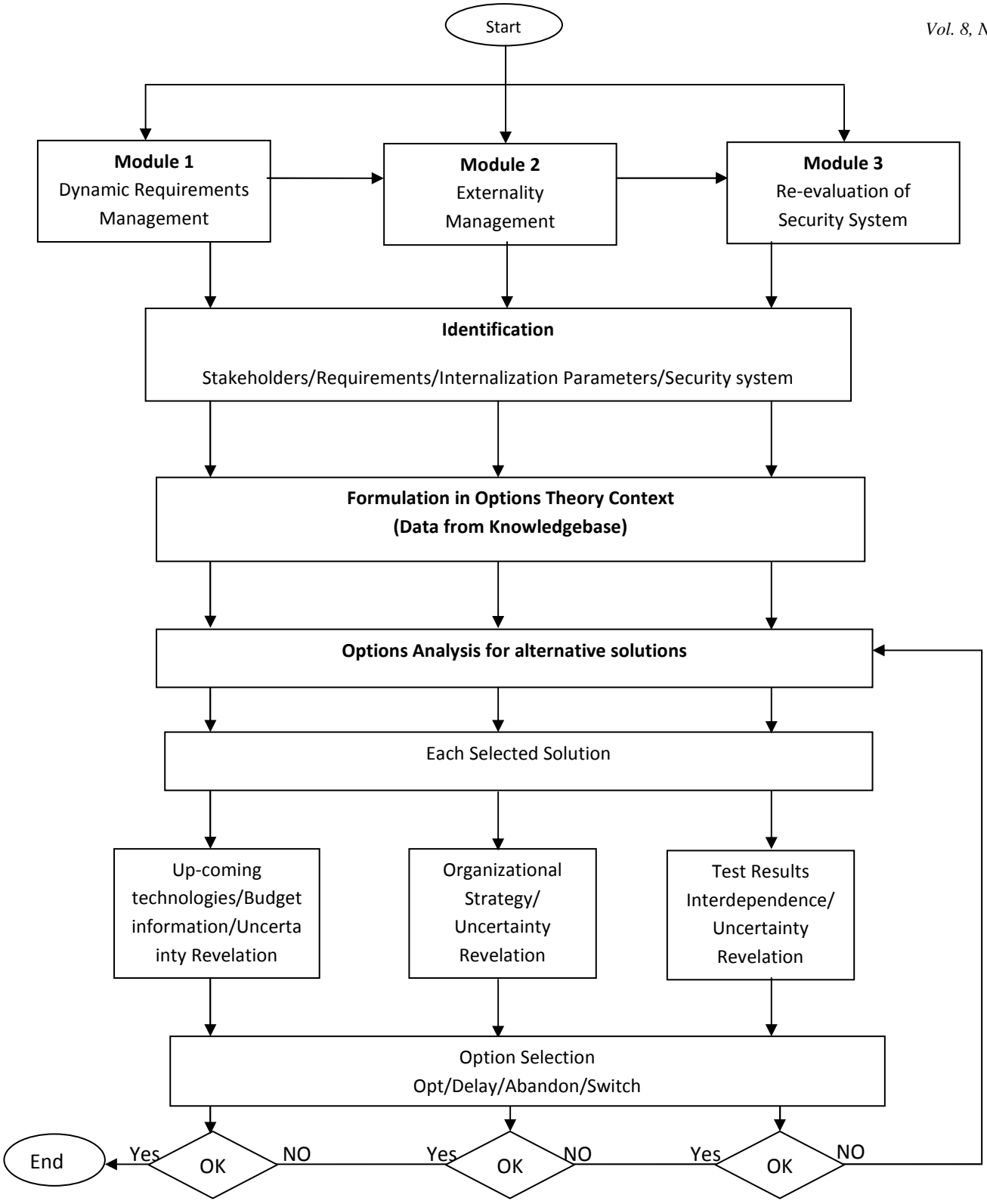


Figure 3. Information flow for Module 1, Module 2 and Module 3

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

i) **Module 1:** Information processing flow of dynamic security requirements management

Dynamically changing security requirement management process starts with the identification of requirement change in an organization. It proceeds with examining all possible solutions for this particular requirement. Each solution is divided into parts and analyzed/compared with system state (determined by the data from knowledgebase), organizational policies/budget and up-coming technologies. The significance of the approach followed by AUMSIS for solution exploration is the options theory. That concurrently analyzes all solutions and decides about each solution to delay, abandon or opt in existing scenario. It provides decision makers analysis reports for the requirement under-consideration, its possible solutions and the pros and cons of each solution according to their organization's information system state.

ii) **Module 2:** Externality management information processing and flow

AUMSIS generates internalization recommendations for the externalities caused by a security system. The security system is already described in knowledgebase according to security mechanisms/services it offers. Internalization process starts with identifying externalities by analyzing system data (from knowledgebase). The next phase is to list all possible solutions (internalization parameters) according to organizational policies and available budget/resources. Each solution is then divided into parts and analyzed using options technique to build organization's internalization strategy considering current system state and organization's future plans. AUMSIS generates internalization results for each internalization parameter to delay, opt and abandon according to existing scenarios as depicted in Figure 3.

iii) **Module 3**: Re-evaluation of security services /mechanisms information processing and flow

AUMSIS helps to build re-evaluation strategy for IT security services/mechanisms considering the uncertain factors i.e. requirements/polices change, vulnerability appearance and interoperability issues that adversely impact evaluation process. The process starts with identifying the boundaries of system for evaluation. It could be the newly adapted solutions in a security system or already evaluated components that need to be re-evaluated. Tests are classified according to the nature of the system under consideration. Uncertainty issues during re-evaluation are dealt using options technique in AUMSIS as tests are prioritized based on pervious evaluation results and vulnerability reports from knowledgebase.

VI. HETEROGENEOUS INFORMATION ISSUE

As the AUMSIS retrieves information from various information sources (i.e. knowledgebase, Internet, organization's policies database) and therefore varies in their structure, syntax and semantics. It is not directly comprehendible by the Option Analyst Agent (OAA). Therefore it is desirable to store information in uniformly accessible and extractable manner. Without considering the operating systems used and the hardware running these softwares. To overcome the issue of heterogeneous information retrieval we have proposed the use of ontologies [14][15] that provide a shared conceptualization of a system or domain. The language used will be Web Ontology Language (OWL) for the development of ontologies. Which is based on strong constructs of description logic and is thus useful to represent any set of rules that are options concepts, organizational policies, internalization parameters etc in case of AUMSIS.

With the help of the options analyst agent these ontologies can be traversed to find the useful information models and to resolve the semantic heterogeneity issues in AUSMIS components. These issues are raised due to the merger of information from various domains i.e. policy database, technological information and vulnerability/malfunctioning reports. It is worth mentioning here that the knowledgebase contains all the organizational polices and rules. This information plays a key role when OAA accesses information from various information sources and formalizes decisions/strategy. Figure 4 depicts the heterogeneous information retrieval framework as follows:

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

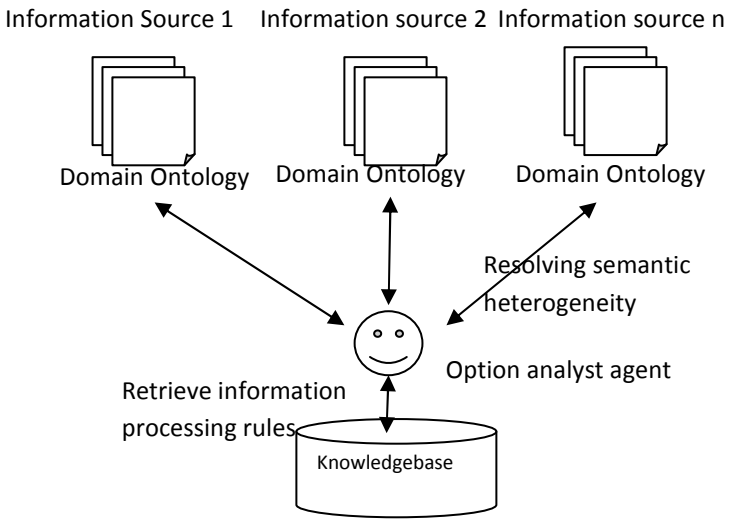


Figure 4. Option analyst agent's communication with ontologies and knowledgebase

VII. AUMSIS ANALYSIS AND VALIDATION

AUMSIS architecture presented in previous sections is based on an in-depth study of its methodological details and manual deployment to SHS and ESAM system in past [2][3][4]. The current AUMSIS design/information flow is about the automated version of options technique's concept for uncertainty management in information security. The architecture currently addresses three main uncertainty issues but is flexible to opt any other problem's mechanism for uncertainty management in information security. Example given below presents the SHS uncertainty management process using AUMSIS in a nutshell. It is worth mentioning here that AUMSIS will be deployed into the organization that interacts and extracts required information about the target system i.e. SHS in this example.

A) Changed requirement request

The process is initiated when a change requirement is identified for SHS system. This could be initiated by an internal source (stakeholders, management and implicit system's request) or by some external source (government enforcing authority) to adapt new standards. Once a change request is identified; it acts as a stimulus to AUMSIS process for the SHS system.

i) Data collection

AUMSIS decides about the optimal solution for a requirement change based on actual system state and within existing circumstances. Knowledgebase provides data to determine system state using vulnerability/malfunctioning reports and system exploitation history with respect to affected SHS services. Dynamic factors like uncertainty revelation, budget information and up-coming technology information are also continuously accessed/considered in solutions formulation process as described in next section.

ii) Options analysis

OAA retrieves information about input data of requirement change for SHS and lists all available solutions for the certain requirement. Each solution is then assigned priorities determined by the up-coming technology, budget information and uncertainty involved in current state. Options theory provides various alternatives to opt, delay or abandon a solution based on uncertainty revelation; also during solution formulation process. AUMSIS analyzes each possible solution by staging its deployment process and wait for additional information that becomes available with the time during exploration and analysis. This additional information normally requires altering the requirement selection strategy; and this facility is factored in as a core feature of AUMSIS. Thus it provides optimal solution about a requirement considering all possible factors that cause uncertainty in determining a solution. The output information of a solution evolution process is stored in knowledgebase that can be used later and provides guidance to examine future strategy.

The newly opted solutions for SHS from a requirement management process may cause positive or negative effects for other interacting partners. Next phase elaborates how these effects are addressed as externalities in AUMSIS.

B) Externality management process

Externalities are the effects borne by the systems that are not involved in a direct communication with the SHS security system. These effects could be positive (that might bring in benefits) or negative (that might cause vulnerabilities) and may appear anytime during the life cycle of SHS system. AUMSIS initiates externality management process by

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

specifying internalization parameters that describes solutions in case of externality occurrences due to SHS system. Organizations (responsible for controlling security system) specify possible internalization parameters according to their security objectives and are stored in knowledgebase.

i) Options Analysis

When an externality is reported/detected for SHS either positive or negative, OAA lists possible internalization parameters for each externality and compares with organizational constraints, which include budget information and organizational policies. These factors are uncertain that may change and affect externality management process. It is also uncertain if a solution will work appropriately. OAA stages each solution into sections and analyzes them individually. All solutions for the externalities of SHS are decided using various options to delay, abandon or alter decision with respect to uncertainty revelation, rational analysis, budget and organizational policies. These factors can be determined using the data from knowledgebase. AUMSIS mechanism of externality management helps to deterministically consider variable factors and to respond accordingly for a specific scenario.

System up-gradation in case of newly installed services for requirement management or externalities solutions (that recommended modifications) requires to re-evaluate the security system to test individual functionality and as a whole interoperability. This factor is also addressed in AUMSIS as a part of a complete solution and described in following section:

C) Initiation of Re-evaluation process

Re-evaluation is performed particularly when new solutions are devised. For example in case of SHS system when the existing system was reconfigured/modified. It is also recommended as periodically scheduled analysis for the complete system. AUMSIS classifies evaluation tests into two major categories i.e. assurance and criticality [7]. Assurance class contains tests to validate performance, serviceability and functionality. Criticality class contains tests that may alter testing strategy and are directly affected by uncertain outcomes/uncertainty issues those are interoperability, technological innovation and budget.

i) Options Analysis

OAA customizes and organizes evaluation strategy for a particular service of SHS based on its history of service failure, vulnerability reports and exploitation history. The information is extracted from the stored data from knowledgebase; which becomes readily available to OAA. Tests are prioritized based on this information and system state. AUMSIS using options theory; provides a deterministic approach to generate evaluation strategy and the ability to alter the evaluation directions. It helps to avoid unnecessary tests that can be determined by the information from uncertain outcomes and uncertainty revelations.

VIII. Conclusion & Future Work

Organizations need to overcome uncertainty issues in their information security management progress due to obvious fact of rapid technological development. They continuously require significant changes in their existing security infrastructure to meet the organizational security objectives and security standards. Organizations also have to invest huge resources and have to go through a cumbersome procedure to keep their system up-to-date. The paper introduced AUMSIS, the infrastructure of an automated system for uncertainty management issues at organizational level based on an in-depth study and manual validation of these concepts in past. The system is capable of managing dynamic issues using options theory mechanism from corporate finance that helps to generate appropriate strategies according to system state. The paper presented the architectural details and information flow for AUMSIS system and its various components. The future intention of this research is the deployment of AUMSIS framework into a software architecture style.

References

[1] Richard Cornes, Todd Sandler, "The Theory of Externalities, Public Goods and Club Goods", Cambridge University Press, June 1996

[2] Ann Cavoukian, "Privacy as a Negative Externality The Solution: Privacy by Design" Workshop on the Economics of Information Security, London, June 24, 2009

[3] Abbas Haider, Yngström Louise and Hemani Ahmed, "Empowering Security Evaluation of IT Products with Options Theory", in 30th IEEE Symposium on Security and Privacy 2009, Oakland, California, USA

[4] Abbas Haider, Magnusson Christer, Yngström Louise and Hemani Ahmed, "A Structured Approach for Internalizing Externalities Caused by IT Security Mechanisms", In Proceesdings of IEEE International Workshop on Education Technology and Computer Science (ETCS 2010), March 2010, Wuhan, China

[5] Abbas Haider, Yngström Louise and Hemani Ahmed, "Option Based Evaluation: Security Evaluation of IT Products Based on Options Theory", In Proceddings of IEEE Eastern European Regional Conference on the Engineering of Computer Based Systems 2009, Novi Sad, Serbia, Pages.134-141

[6] J. Mun, "Real Options Analysis - Tools and Techniques for Valuing Strategic Investments and decisions", Wiley, Finance, 2002

[7] Abbas Haider, Yngström Louise and Hemani Ahmed, (2009), "Adaptability Model Development for IT Security Evaluation Based On Options Theory" in proceedings of IEEE/ACM 2nd International Conference on Security of Information and Networks (SIN 2009), North Cyprus

[8] Abbas Haider, Raza Asad , Louise Yngström, Ahmed Hemani, "Evaluation of ESAM using Architectural Tradeoff Analysis Method", Project Report -VERVA, December 2008

[9] Kurt Helenelund, Stephan Urdell, Bo Sehlberg, Anders Bremsjö, Anders Lindgren, Jan Lundh, Christer Marklund, "SHS Version 1.2 Protocols", VERVA - Swedish Administrative Development Agency, 2007

[10] Wilson, D. ; Greig, A. ; Gilby, J. ; Smith, R., "Intelligent automated inspection, representing the uncertainty of the real world", IEE Colloquium on Intelligent Sensors (Digest No: 1996/261), 19 Sep 1996, pages 11/1 - 11/6

[11] McVicker, M.; Avellino, P.; Rowe, N.C., "Automated Retrieval of Security Statistics from the World Wide Web" IEEE SMC Information Assurance and Security Workshop, 2007, 20-22 June 2007, pages 355 - 356

[12] Abbas Haider, Yngström Louise and Hemani Ahmed, "Security Evaluation of IT Products: Bridging the Gap between Common Criteria (CC) and Real Option Thinking" in proceedings of World Congress on Engineering and Computer Science (WCECS 2008), 22-24 October, 2008, San Francisco, USA

[13] Nick Jennings, Michael Wooldridge, "Software Agents", IEE Review, January 1996, pp 17-20

[14] Thomas R. Gruber: Automatically Integrating Heterogeneous Ontologies from Structured Web Pages. Int. J. Semantic Web Inf. Syst. 3(1): 1-11 (2007)

[15] Xiaomeng Su, Mihhail Matskin and Jinghai Rao. "Implementing Explanation Ontology for Agent System". In Proceedings of the 2003 IEEE/WIC International Conference on Web Intelligence, WI'2003, Halifax, Canada, October, 2003. IEEE Computer Society Press

AUTHORS PROFILE

Haider Abbas has been working as doctoral student at Department of Electronic Systems, KTH, Sweden. Mr. Abbas has authored more than 10 international publications and has been working for various governmental and private projects in Pakistan and Sweden.

Christer Magnusson is Senior Lecture at the Department of Computer and Systems Sciences, Stockholm University, specialized in IS/IT Security and IS/IT Risk Management. Before joining SecLab, Christer was Head of Corporate Security and Risk Manager at Sweden Post and CEO of Sweden Post Insurance AB and Sweden Post Reinsurance S.A. He has also held the position as Head of Corporate Security in the Ericsson group. In 1999, Christer was awarded the SIG Security Award by the Swedish Computer Society and in 2000 the Security Award by the Confederation of Swedish Enterprise as recognition of the models and the integrated processes regarding IS/IT Risk Management, that he developed as a part of his research studies. Dr. Christer is a member of the advisory committee of the Swedish Emergency Management Agency. He serves on the risk and security board of the Confederation of Swedish Enterprise (NSD). He is also an adviser in Corporate Governance, Compliance, Risk Management, and Information and ICT Security to government agencies as well as trade and industry.

Louise Yngström is a professor in Computer and Systems Sciences with specialization in Security Informatics in the department of Computer and Systems Sciences at Stockholm University. Her research base is Systems Science which she since 1985 has applied within the area of ICT security forming holistic approaches. Her research focuses various aspects on how ICT security can be understood and thus managed by people in organizations, but also generally on criteria for control. She has been engaged in various activities of the International Federation of Information Processing, IFIP, since 1973; the Technical Committee 3(TC3) with an educational scope, the TC9 with focus on social accountabilities of ICT structures and the TC11 with focus on ICT security. She founded the biannual conference WISE (World Conference on Security Education) in 1999. She was engaged in European networking for curricula developments within ICT security and the Secured Electronic Information in Society working for e-Identities during the 1990's. Since 2000 Dr. Louise is involved in introducing ICT security in academic and business life in African countries through her research students who simultaneously with their research are academic teachers in their home countries. Over the years she has traveled and

networked extensively with international peers. Presently she is the principal advisor of seven PhD students.

Ahmed Hemani has been working as professor and head of post-graduate studies at Dept. of ES, School of ICT, KTH, Sweden. Dr. Hemani has authored more than 100 international publications. He is participating in and leading some national and EU projects.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Driving Architectural Design through Bussiness Goals

Lena Khaled
Software Engineering Department
Zarqa Private University
Amman, Jordon
lenaumleen@yahoo.com

***Abstract* - Architectural design must encompass changes to business goals, their relations to quality attributes overtime and their results upon building the final specific systems. This paper discusses the effect of business goals on building the architectural design on any system. It describes the relationship between business goals and system qualities, and how these qualities are met during architectural design. This paper also describes how the qualities have an effect on the decisions of building the architectural design on any system. The role of the agent is described through the process of building.**

Keywords- Bussiness Goals; Qualities; Architectural Design Decisions.

I. INTRODUCTION

An important part in the design phase and construction of any complex system is its architecture. A good architecture is that system which satisfies key requirements such as performance, reliability, portability, interobelability, availability, security, safety, scalability and other attributes. A bad architecture causes a lot of damage. According to IEEE, "*Architectural design is the fundamental part in any system which includes components of a system, their relationships to each other and to the environment and the basic principles which guide the design and evolution of that system*." [5, 7]

On the other hand, business goals are the foundation on which any architectural design is built upon. Architectural design is built to achieve mission goals. Business goals come in many types and at many levels of abstraction; therefore, architectural design is a bridge between business goals and the achieved system.

This paper describes how business goals cause the progress to build architectural design to any system; many processes need to be applied in order to reach the final system. This paper is organized as the following: section 3 defines business goals and attributes and then it defines the relation between them. Section 4 describes the architectural design decisions and its major phases when defining the design process. Section 5 defines the main result of this paper; it talks about how the final architecture is built, the role of software agent through building and software quality metrics. You can find the conclusion in section 6.

II. RELATED WORKS

The first piece of related works is reported in the most depth in Perry. Perry Dewayne and Wolf Alexander worked on studying the foundations of software architecture, they developed an intuition for software architecture by applying to several disciplines of architecture, and they presented a model of software architecture which consists of three components: elements, form and rationale [12].

The second related works is [5]. Garlan examined important trends of software architecture in research and practice, and speculated on the important emerging trends: aspiration and challenges.

The third paper in related works is [1]. Ozkaya, Kazman and Klein worked on a case that represented the architectural patterns which carries the economic value in the form of real options. They summarized their observations in evaluating the relative value of patterns using real option value models on a model problem and they looked carefully on how economics-informed approaches can provide better insights for the selection of situated design strategies.

The fourth paper in related works is [1]. This paper describes how architectural decisions are made from quality attributes. It presents a set of steps that allowed moving from quality attribute requirement to design fragment based on achieving that requirement. All of these are demonstrated through an application of an embedded system as in example.

This paper works on the aim of business goals and their relationships to the quality attributes, and how this relation can drive the architectural design to any complex system.

III. BUSINESS GOALS AND QUALITIES

Understanding bussiness goals and their relations to the qualities is a critical part of building the architectural design of any system; we cannot easily use architectural design or other solutions without understanding the concepts of both bussiness goals and qualities. Therefore, these goals and qualities drive the architectural design of the system [3].

Driving architectural design through goals need an early method used to generate and refine qualities, which is called QAW (Quality Attribute Workshop). This gets qualities that are mapped to business goals scenarios for the qualities which are built by stakeholders according to the main goals. All these scenarios specify whether a system satisfies the

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

user's requirements or not [3, 2]. The quality attributes must be well understood and expressed early in the development of the life cycle of the system, so the architect can design an architecture that will satisfy these qualities. QAW is a method to discover, document, prioritized the system's quality attributes early in its life cycle [2].

A. *Bussiness goals*

Business goals are the parts that drive the methods of design, and are the elements that shape the architecture. The important thing is that all business goals that correspond to quality attributes will view the end of the system [3, 6].

According to [10], we can define a goal as a state of events in the world that users would like to achieve. Goals can be rather business goals or system goals. Business goals are states that an individual or an organization wants to reach.

B. *Quality attributes*

In manufacturing, the concept of the quality means that the product should meet its requirements, but the popular vision of the quality is that it is an intangible attribute. Terms of bad or good quality represent how people talk about something vague which they don't propose to define. Quality attributes describe the property of the system that refers to its fitness for use. The term, non-functional requirement, is a synonym for quality attribute or attribute. [13, 9, 4]

The international standard on software product qualities classifies software quality as six main attributes: functionality, reliability, usability, efficiency, maintainability, and portability. Despite the fact that there are many quality attributes, reliability and maintainability are the main quality criterions and many of these attributes are created at business levels and are better viewed as business goals [6, 8]. Figure 1 illustrates how goals and attributes affect each other.

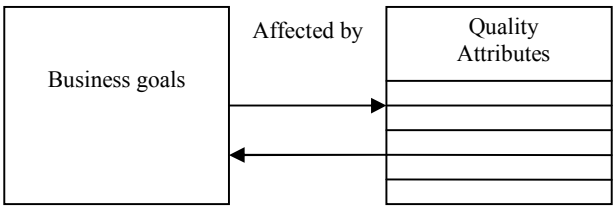


Figure 1. The relation between goals and attributes

IV. ARCHITECTURAL DESIGN DESCISIONS

Architectural design decisions play an important position in designing the architecture for any system. We can define an architectural design decision as: a description of the set of architectural modifications to the software, the principles and rules to the design, design constraints and additional requirements that realize requirements on a given architecture. During the process of architectural design, the software designer has to make a number of decisions that affect the system and its development process [13].

Figure 1 shows how the decisions of a designer affect the design process. According to the experience of the designer, the main decisions are the answer to the following questions [13]: What is the main approach that has been used to structure the system? What is the strategy that has been used to control the operations in the system? How is distribution across the occurring system? What is the appropriate style for the system? How do the evaluations of the architectural design occur? How should the architecture of the system be documented?

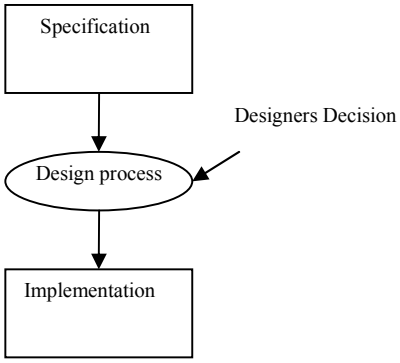


Figure 2. The role of designer's decisions on a design process

When developing systems, the process of the design is divided into two distinct phases. Figure 3 describes these phases [14].

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

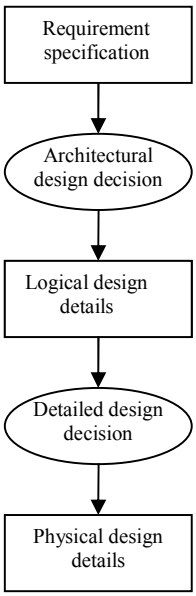


Figure 3. The major phases of the software design process

1) In the first phase, the designer develops a high level of solutions (the architectural or logical design) in which only the abstraction view of the model is described without details.Here, the system is described as black-box system.

2) In the second phase, the abstract view is going into details(the details or physical design) . Turning the black box into white box , the output from this phase provides the specifications for the programmer.

Making a decision to any architecture is based on the relation between goals and qualities.

V. FINALIZING ARCHITECTURAL DESIGN

The architectural design of any system can be defined as the structure of the system, which consists of software components, the external properties of these component and their relationships. The architectural design affects the performance, maintainability, robustness. The particular style and structure chosen for an application may therefore depend on non functional requirements [14, 13].

A. The role of the agent

Software agents are autonomous software entities that navigate through environments and can either work alone or with other agents to achieve the goal.

The software agent plays an important role in the cycle that is represented in figure 3. The fundamental reason for using an agent to build the architectural design for any system is based on the concept that many users (stakeholders) within an organization have many different business goals. Therefore, by linking software agent to this cycle, it becomes possible to give the main impact for choosing a specific architectural design decisions that arise through building the architecture [6]. The architect needs many decisions to build the final system. Then, the software agent helps the architect to choose the main decisions that make the architect achieving his goal to meet the user requirements.

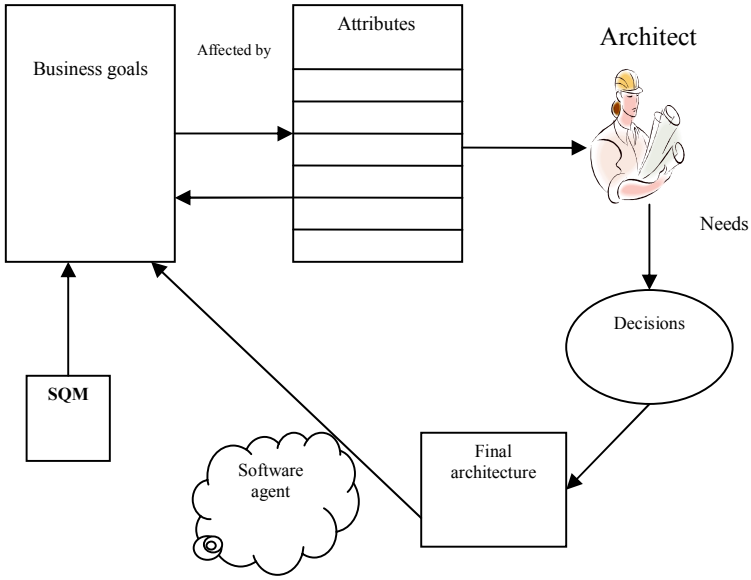


Figure 4. Driving Achitectural Design through bussiness goals

Figure 4 represents the result of this paper; it describes how business goals are affected by the attributes and how this relation affects on the decision of the architect and in which from these decisions the final architecture is built. The role of the agent is described through the figure. The agent is especially raised when we want to move from old to new goals.

Creating goals is an important thing to the products and process. These goals should be considered by the architect through mechanisms, which are driven by models of business. There are many types of mechanisms such as the SQM that appears in figure 4 and is defined as a mechanism for measuring goals. Measurements enable the organization to make its process better than before, in addition to the controlling of the software project and assessing the software quality. Other measurement can be used instead of SQM such as the Quality Function Deployment approach (QFD), the Goal/Question/Metric paradigm (GQM).

Management is done better by measurement, measurement enables the organization to improve the

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

activities of the quality management: planning, control and asses the quality of the software that produced.

The quality management activities check the project output to ensure that they are dependable with the organizational standards and goals. The development team should be independent from the quality assurance team so that they can take a goal view of the software. An independent team should be responsible for managing the quality to maintain the project schedule and budget. As a result, this team should not be associated with any particular development group. An independent team ensures that the organizational goals of quality are not compromised by schedule consideration and short-term budget [13].

The evaluation of the overall work will reflect its relationship to the quality of work that is expected from decisions which were made by the architect before. It is defined as the process of verifying satisfaction of requirements. The final architectural design is very difficult to be evaluated because the correct test of the architecture is in how it meets the functional and non-functional requirements after it has been deployed [13].

VI. CONCLUSION

Architectural design is one of the most important phases in the life cycle of any system. Building an architectural design for any system to any organization must include the changes that occurred to business goals. This paper presents how an architectural design was drove from business goals and this is done through explaining the relation between business goals and non-functional requirements (also called quality attributes or attributes). It describes how an architect needs decisions to build the architecture and how these decisions are built upon the previous relation. The role of an agent is shown through this paper to show the movement from old business goals to new business goals.

REFERENCES

[1] Bachman F,Bass l,klein M,"Moving From Quality Attribute Requirment to Architectural Decisions,"Software Engineering Institute,USA.

[2] Barbacci M,Ellison R,Lattanze A,Stafford J,Weinstock C,and Wood W,"Quality Attribute Workshops(QAW),technical report CMU/SEI,2003.

[3] Bass L,Clements P, Kazman R,Nord R,"Architectural Business Cycle Revisted," Software Engineering Institude,Caregie Mellon,2009.

[4] Berenbach B,Paulish D,Kazmeier J,Rudorfer A,"Software and systems requirments engineeering in practice,"McGrawHll,2009.

[5] David Garland,“Software Architecture:a Roadmap,” ACM Press,2000.

[6] Gross D, Yu E,"Evolving System Architecture to Meet Changing Bussiness Goals:an Agent and Goal –Oriented Aapproach,"University of Toronto.

[7] http://www.sei.cmu.edu/ Carnegie Mellon,2007.

[8] Jalote P,"A concise introduction to software engineering," Springer,2008.

[9] Kan S,"Metrics and models in software quality engineering,"Addison Wesely,2003.

[10] Liu l,yu eric,"From Requirment to Architectural Design – using Goals and Scenarios," University of Toronto.

[11] Ozkaya I, Kazman R, Klein M,"Quality attribute based economic valuation of architectural patters,"SEI,9TH international conference on software engineering,IEEE,2007.

[12] Perry D.,Wolfs A.,“Foundations of the study of software architecture,” ACM Sigsoft, vlo.17, no. 4, 1992.

[13] Sommerville I,"Software Engineering,"Pearson education, 2007.

[14] Widhani A,Boge S,Bartelt A,Lamersdof W,"Software architectural and patterns for electronic commerce systems,"University of Hamburg,2002.

DISTRIBUTED INFORMATION SHARING COOPERATION IN DYNAMIC CHANNEL ALLOCATION SCHEME

Mr.P.Jesu Jayarin,
Research Scholar,
Sathyabama University,
Chennai-119, India.
jjayarin@gmail.com,

Dr.T.Ravi,
Prof&Head ,Dept of CSE,
KCG college of Technology
Chennai-97, India.

Abstract

The Channel allocation method is an Cooperative Asynchronous multichannel MAC which introduces an Distributed Information Sharing (DISH) to be used in a distributed flavor of control-plane cooperation, as a new approach to wireless protocol design and then apply it to multichannel medium access control(MAC) to solve the MCC problem. The basic idea is to allow nodes to share control information with each other such that nodes can make more informed decisions in communication. Medium access control (MAC) protocols play a major role to create a wireless communication infrastructure. In Wireless network, transmitter-receiver pairs make independent decisions, which are often suboptimal due to insufficient knowledge about the communication environment. So, the new concept DISH is introduced and overcomes the problem occurred in the MAC protocol. The DISH concept avoids collision and re-transmission among nodes. The notion of control-plane cooperation augments the conventional understanding of cooperation, which sits at the data plane as a data relaying mechanism. In a multichannel network, DISH allows neighboring nodes to notify transmitter-receiver pairs of channel conflicts and deaf terminals to prevent collisions and retransmissions. Based on this, we design a single-radio cooperative asynchronous multichannel MAC protocol called CAM-MAC. When the CAM-MAC is used in illustration purposes, we choose a specific set of parameters for CAM-MAC. First we analyze the throughput to 91% of the system bandwidth to 96%, then saturate 15 channels and compare the result, this provides an good result in implementing hardware.

Keywords-*Distributed information sharing (DISH), control-plane cooperation, CAM-MAC, multichannel coordination problem, MAC protocol, ad hoc networks.*

1. INTRODUCTION

The Cooperative asynchronous multichannel MAC is to allow nodes to share control information with each other such that nodes can make more informed decisions in communication. This notion of control-plane cooperation augments the conventional understanding of cooperation, which sits at the data plane as a mechanism for intermediate nodes to help relay data for source-destination pairs Asynchronous Multichannel MAC is to allow nodes to share control information with each other such that nodes can make more informed decisions in communication.

The new approach DISH is implemented first and then medium access control to solve MCC problem. In a multichannel network, DISH allows neighboring nodes to notify transmitter-receiver pairs of channel conflicts and deaf terminals to prevent collisions and retransmissions. Based on this, we design a single-radio cooperative asynchronous multichannel MAC protocol called CAM-MAC. This is used to provide a good result in implementation. The heartily focused on giving a distributed solution in minimal cost, efficient and sufficient solution in the real world application is to allow the nodes to share the control information with all the other nodes in communication. Example for instance, transfers of data from one node to another node in a distributed environment without collision and re-transfer of data. DISH enables nodes to store channel usage information at their neighbors, and retrieve this information whenever it is needed. It is not a compulsory coordination mechanism but it provides an cooperation with other nodes and operates well at all condition. In this network does not rely on coordination and performs well. Most wireless LANs are single channel systems.

However, as the number of nodes communicating increases, systems with a single channel suffer declining performance. Contributing to the problem are the well-known

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

hidden and exposed terminal problems. To combat these problems there is growing interest in multi-channel systems. Indeed, the IEEE 802.11 standard already has multiple channels available for use. The IEEE 802.11a physical layer has 12 channels,8 for indoor and 4 for outdoor use. IEEE 802.11b has 14 channels, 5*MHz* a part in frequency. To avoid channel overlap, the channels should have at least 30*MHz* guard bands; typically, channels 1, 6 and 11 are used for communication. In a multi-channel system, the transmitter and receiver must both use an agreed upon channel for communication. This introduces a channel coordination problem. As well, the hidden and exposed terminal problems remain in the multichannel setting.

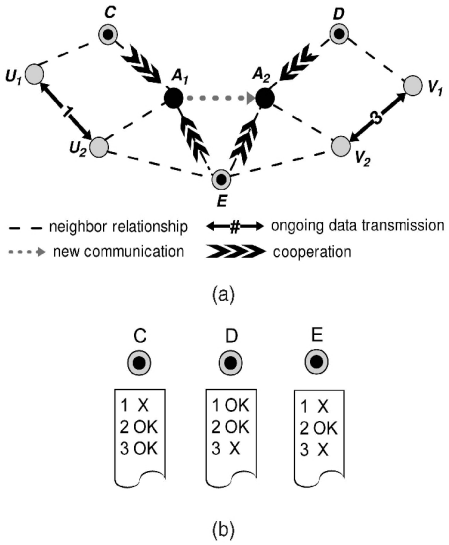


Fig. 1. An illustration of the DISH idea. (a) A multichannel scenario.

(b) Knowledge at individual nodes. By consolidating the knowledge at nodes C and D, or acquiring knowledge from node E, it shows that the conflict-free channel is channel 2.

Based on the idea of DISH, we design a single-radio cooperative asynchronous multichannel MAC protocol called CAM-MAC for ad hoc networks. We evaluate CAM-MAC from both theoretical and practical perspectives, where we choose a specific set of protocol parameters for illustration and evaluation purposes:

1.We show that its throughput upper bound is 91 percent of the system bandwidth and its average throughput approaches this upper bound with a mere gap of 4 percent,

2.We show that it can saturate 15 channels at maximum and 14.2 channels on average, which indicates that, although CAM-MAC uses a control channel, it does not realistically suffer from control channel bottleneck,

3.To investigate the value of cooperation we compare MAC and observe a throughput ratio of 2.81 and 1.70 between them in single-hop and multihop networks, respectively,

4.We compare CAM-MAC with three recent and representative multichannel MAC protocols, MMAC, SSCH, and AMCP , and the results show that CAM-MAC substantially outperforms all of them.

For a further and more realistic validation, we implemented CAM-MAC on COTS hardware and conducted experiments. To the best of our knowledge, these prototypes are the first full implementation of single-radio asynchronous multichannel MAC protocols.

We review literature in Section 2, and identify new challenges to designing a cooperative protocol in Section 3. Then, we present the protocol details in Section 4. Following that, Section 5 provides performance results in various scenarios, and Section 6 describes our implementation and experiments. and, finally, give concluding remarks in Section 7.

2. MULTICHANNEL MAC PROTOCOLS

Multichannel MAC protocols for ad hoc networks can be categorized into two general classes as below.

2.1 Single-Radio Solutions

2.1.1 Control-Data Window Schemes

MMAC assumes the IEEE 802.11 power saving mode and divides time into beacon intervals. Each beacon interval is 100 ms and consists of a 20-ms ATIM window and an 80-ms data window. Nodes negotiate and reserve channels in the ATIM window on a common channel, and transmit data in the data window on multiple channels concurrently. The data window size is fixed, and hence, it has to be set according to the maximum data packet size, leading to inefficiency. By contrast, MAP varies the data window size and avoids this problem. However, like MMAC, its reservation interval (i.e., control window) is still fixed, and thus, both protocols suffer from the inflexibility to different node densities: at low density, the control window has long idle time; at high density, the control window cannot

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

accommodate all negotiations and some nodes have to wait for the next control window. Further-more, MMAC and MAP requires time synchronization over the entire network, which is a notoriously hard task involving considerable overhead and complexity, and compromises scalability. LCM MAC, on the other hand, allows each neighborhood to negotiate the boundaries of control-data windows independently, in order to avoid time synchronization. However, the negotiated window size can hardly fit for all nodes in the neighbor-hood, and this window negotiation, plus a fine-tuning mechanism, considerably complicates the protocol. Besides, it has a starvation problem lacking in an appropriate solution. Finally, all these control-data window schemes have a common problem: during each control window, all channels other than the common channel cannot be utilized, resulting in channel underutilization.

2.1.2 Channel Hopping Schemes

In SSCH , each node hops among all channels according to a pseudorandom sequence such that neighboring nodes will have channels overlap in time periodically. Since a trans-mitter can only communicate to a receiver when they hop. In CAM-MAC, each node stays on a common channel and only switches channel when a data exchange is established successfully or finished. This avoids switching channel too often and, due to the common channel, does not incur large delay. Besides, again, no clock synchronization is required.

2.1.3 Routing and Channel Assignment Schemes

CBCA combines channel assignment with routing. It proposes to assign each set of intersected flows, called a component, with a single channel, in order to avoid channel switching delay, node synchronization, and scheduling overhead at flow-intersecting nodes. CAM-MAC uses a control channel, which automatically avoids the problem of node synchronization and scheduling overhead. Regarding channel switching delay, its effect on network performance is much less than MCC problems: the channel switching delay is 40-150 μs but a channel conflict can collide at least one data packet whose delivery several and even tens of milliseconds.

In fact, CBCA shifts complexity from the MAC layer to the routing layer. Also, compared to packet, link, and flow-based channel assignments, it has the least flexibility in exploiting multichannel diversity. Each component, which spans all intersecting flows, can only use one channel. As a consequence, any two nodes in a common component cannot transmit simultaneously unless they are at least three or four hops apart (depending on the interference range). In a single-hop network, since all flows are intersected, a multichannel network degrades to a single-channel network.

2.2 Multiradio Solutions

Using multiple radios can easily solve MCC problems by dedicating one radio for monitoring channel usage information.DCA uses two transceivers, one for exchanging control packets and the other for exchanging data packets. The control packets are used to allocate the channels on the data transceiver on demand. A multichannel CSMA protocol assumes the number of channels to be equal to the number of transceivers per node, so that all channels can be used simultaneously. This is very expensive. A protocol similar to DCA in that it also dedicates a transceiver for control purposes, but the difference is that channel selection is done at the receiver end based on signal-to-noise ratio. MUP also uses two transceivers but it allows both transceivers to exchange control messages and data packets. xRDT extends RDT, which uses a (possibly different) quiescent channel for each node to receive packets, by adding a busy-tone radio to each node in order to inform the neighborhood of ongoing data reception, in order to avoid collision and deafness. It proposed link-layer protocols for routing in multiradio multichannel ad hoc networks. Each node is assigned a fixed interface for receiving packets and multiple switchable interfaces for transmitting packets. This is similar to the idea of quiescent channel but uses more radios to simplify overcoming MCC problems. Obviously, the key drawback of multiradio protocols is the increase of device size, cost, and potentially energy consumption.

3. CAVEATS TO COOPERATIVE PROTOCOL DESIGN

We identify three major issues in designing a cooperative MAC protocol, which will adversely affect protocol performance unless properly addressed.

3.1 Control Channel Bottleneck

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Using a dedicated control channel can facilitate the design of a cooperative protocol, because a control channel provides a unique rendezvous for nodes to disseminate, gather, and share information. However, this design scheme may come with a drawback: When a large number of channels and nodes are present, the single control channel which is used to set up communications can be highly congested and become a performance bottleneck.

3.2 Cooperation Coordination

An MCC problem can be identified by multiple neighboring nodes, and hence, their simultaneous response of sending cooperative messages will result in collision. This creates an issue of cooperation coordination. One solution is to make neighbors sequentially respond via a priority-based or slot-based mechanism, thereby ensuring all cooperative messages to be transmitted without collision. However, this is very inefficient because 1) there can be many wasted (idle) intervals because not all neighbors may identify the problem and 2) cooperative messages pertaining to the same MCC problem carry redundant information, and hence, receiving all of them is not necessary. Another solution is to let each neighbor send such messages probabilistically, in order to reduce the chance of collision. However, an optimal probability (optimal in the sense of minimizing the chance of collision) is hard to determine, and such a scheme can result in no response which essentially removes cooperation. Therefore, a simple yet effective coordination mechanism is needed.

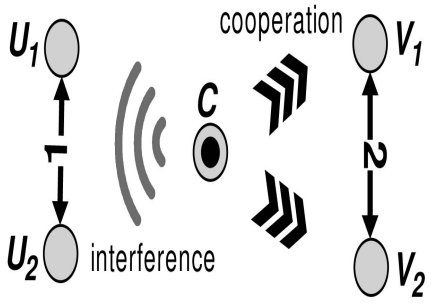


Fig. 2. Cooperation interference

3.3 Cooperation Interference

This issue means that the cooperative messages sent by neighbors for a transmit-receive pair can unconsciously cause interference to another (nearby) transmit-receive pair. This is a new type of interference created by the introduction of cooperation, and our simulations found that it frequently happens and considerably intensifies channel contention. As such, a mechanism needs to be devised to address this deleterious side effect.

4. PROTOCOL DESIGN AND ANALYSIS

Our assumption is that each node is equipped with a single half-duplex transceiver that can dynamically switch between a set of orthogonal frequency channels but can only use one at a time. We do not assume specific channel selection strategies; CAM-MAC runs on top of any such strategy. For quantitative performance evaluation, we will consider two strategies in simulations and experiments: 1) RAND selection, where a node randomly selects one from a list of channels that it deems free based on its knowledge, and 2) most recently used (MRU) selection, where a node always selects its MRU data channel unless it finds the channel to be occupied by other nodes, in which case RAND selection strategy is used.

We do not assume equal channel bandwidth or channel conditions such as noise levels; these can be taken into account by channel selection strategies (e.g., choose the channel with the highest SNR) which are not in our assumptions. We also do not assume any (regular) radio propagation patterns, nor assume any relationship between communication ranges and interference ranges.Intuitively,none of the nodes is responsible for providing cooperation; a node cooperates if it can (it is idle and overhears a handshake that creates an MCC problem), and simply does not cooperate . Actually, there often exists at least one neighboring node that can cooperate, and even in the worse where no one can cooperate, the protocol still proceeds (as a traditional noncooperative protocol).

4.1 Protocol Design

One channel is designated as the control channel and the other channels are designated as data channels. A transmitter and a receiver perform a handshake on the control channel to set up communication and then switch to their chosen data channel to perform a DATA/ACK

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

handshake, after which they switch back to the control channel.A transmitter sends a PRA and its receiver responds with a PRB, like IEEE 802.11 RTS/CTS for channel reservation. Meanwhile, this PRA/PRB also probes the neighborhood inquiring whether an MCC problem is created (in the case of a deaf terminal problem, it is probed by PRA only). Upon the reception of the PRA or PRB, each neighbor performs a check and, if identifying an MCC problem, sends an INV message to invalidate the handshake (the receiver can also send INV after receiving PRA, since it is also one of the transmitter's neighbors). If no INV is sent and the transmitter correctly receives PRB, it sends a CFA to confirm the validity of PRA to all its neighbors (including the receiver), and the receiver will send a CFB to confirm the validity of the PRB if it correctly receives CFA. This marks the end of a control channel handshake. If any INV is sent, the handshake will not proceed and the transmitter will back off. The NCF is merely used by the transmitter to inform its neighbors that the PRA and CFA are invalid when it fails to receive CFB (the receiver gets INV after sending PRB).

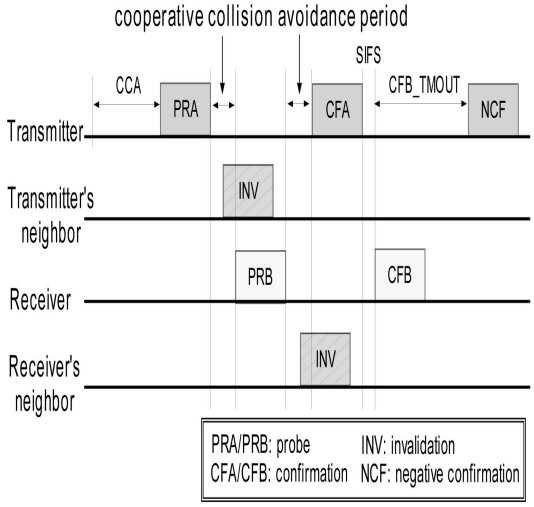


Fig. 3. The CAM-MAC control channel handshake.

The cooperative collision avoidance period is for mitigating INV collision caused by multiple neighbors sending INVs simultaneously. It is a simple CSMA-based mechanism where each neighbor schedules to send INV at a random point in this period and continues sensing the channel. Once the node that schedules at the earliest time starts to send, others in its vicinity cancel sending their INVs (a receiver can also cancel its PRB).

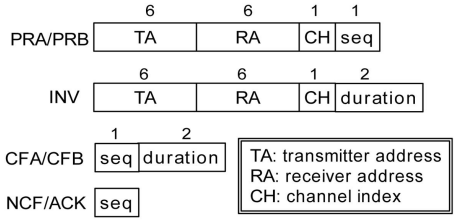


Fig. 4. A possible set of frame formats

Channel usage table, shown in Fig. 5. Note that the until column does not imply clock synchronization: It is calculated by adding the duration in a received CFA/ CFB/INV message to the node's own clock. Similarly, when sending INV, a node does a reverse conversion from until to duration using a subtraction.

TA	RA	CH	until
A_1	A_2	1	11:30:52
B_1	B_2	3	11:30:56

Fig. 5. Channel usage table.

Also note that this table is by caching overheard information while not by sensing data channels. This is because sensing data channels often obtains different channel status at the transmitter and the receiver, and resolving this discrepancy adds protocol complexity. In addition, this may lead to more channel switching's and radio mode (TX/RX/IDLE) changes and thus incurs longer delay.

4.2 Caveats Revisited

Now, we explain how we address the caveats stated in Section 3 in the design of CAM-MAC.

4.2.1 Cooperation Coordination

Recall that this issue is to coordinate multiple neighbors to send cooperative messages as efficiently as possible. We address this using the cooperative collision avoidance period described in Section 4.1. It ensures that only one node will send out a cooperative message (INV) in each single-broad-cast region, assuming that propagation delay is negligible. We found via simulations that this can reduce 70 percent-85

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

percent collisions between cooperative messages.

In case that collisions still happen (due to propagation delay or because not all cooperative nodes can hear each other), it is not a serious problem because CAM-MAC makes such collisions meaningful by using negative feedback only. That is, since INV always means invalidation, a collision resulting from INV still conveys that the hand-shake should not proceed. Actually, using negative feed-back is a logical design. First, a node expects a binary feedback since it selects one channel, instead of selecting a list of channels which needs multibit feedback indicating busy/free channels. Second, sending a positive feedback can be misleading because ensuring no MCC problem requires full information while a cooperative node cannot guarantee to have.

5. PERFORMANCE EVALUATION

We evaluate and compare five protocols, namely IEEE 802.11, CAMMAC-RAND, CAMMAC-MRU, UNCOOP-RAND, and UNCOOP-MRU, using a discrete-event simulator which we developed on Fedora Core 5 with a Linux kernel of version 2.6.9. In these five protocols, IEEE 802.11 is used as a baseline in comparison, X-RAND and X-MRU are two versions of protocol X using RAND and MRU channel selection strategies, respectively. The protocol UNCOOP is identical to CAM-MAC except that the cooperation element is removed, i.e., neighboring nodes do not participate in communication by sending INV messages. This comparison will enable us to investigate the value of cooperation.

We use three performance metrics: 1) aggregate (end-to-end) throughput, 2) data channel conflict rate, defined as the packet collisions on data channels per second over all nodes, and 3) packet delivery ratio, defined as the number of data packets successfully received by destinations. There is one control channel and five data channels with bandwidth 1 Mbps each. PHY and other MAC layer parameters, i.e., PLCP, SIFS, and retry limit, are the same as in IEEE 802.11 . Each source generates data packets with 2-Kbyte payload according to a Poisson point process. The cooperative collision avoidance period is 35μs . In the comparison of CAM-MAC and UNCOOP, we ignore channel switching delay as both protocols use the same handshake. However, in comparison to the other protocols, namely MMAC, SSCH, and AMCP, we use the parameters that they, respectively, use, including channel switching delay. The evaluation is based on 1, Impact of traffic load 2, Impact of data payload size 3, Impact of the number of nodes.

6. IMPLEMENTATION

There are 30 nodes forming 15 disjoint flows in a single-hop network. In the first set of simulations, the flows are always backlogged and the number of channels varies from 2 to 12. We see that CAM-MAC achieves a throughput of 11.86 Mbps while AMCP achieves 8.5 Mbps when there are 12 channels, which indicates a ratio of 1.40. Furthermore, AMCP saturates at 10 channels whereas CAM-MAC still exhibits a growing trend beyond 12 channels. In the second set of simulations, there are four channels and the traffic generation rate varies from 8 Kbps to 8 Mbps.Both CAM-MAC and AMCP have equal throughput at light traffic load, but apparent difference appears at medium to high load, and finally, CAM-MAC saturates at 5 Mbps while AMCP saturates at 4.2 Mbps.

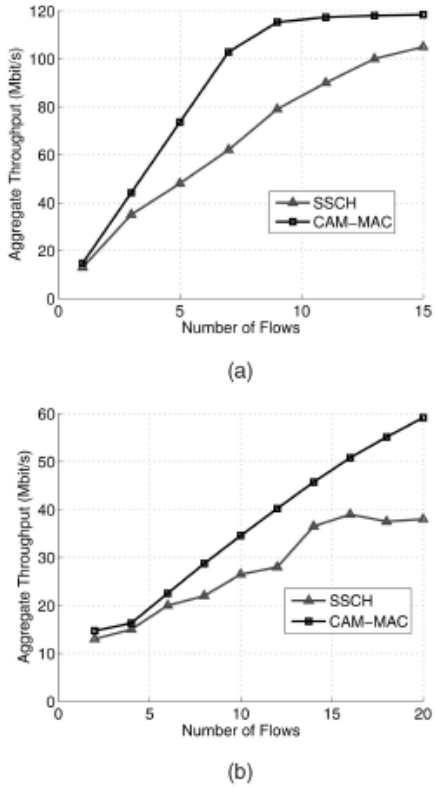


Fig. 6. Comparison with SSCH. (a) Disjoint flows. (b) Nondisjoint flows

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

We implemented CAMMAC-RAND, CAMMAC-MRU, UNCOOP-RAND, and UNCOOP-MRU on COTS hardware implementation of single-radio asynchronous multichannel MAC protocols for ad hoc networks and a multichannel time synchronization protocol. It periodically exchanges beacon packets but data handshaking was not implemented. It provide a test bed for routing and channel assignment via statical manual configuration instead of hardware implementation. In a multichannel MAC protocol, which is designed for sensor network data collection applications assuming the many-to-one traffic pattern.

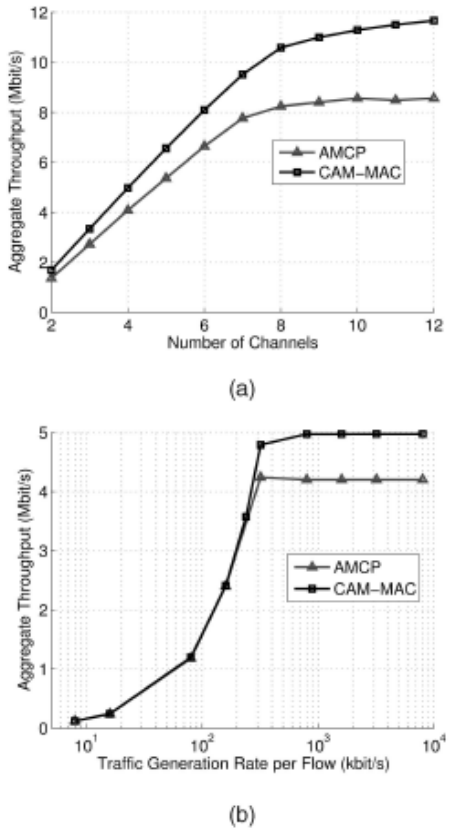


Fig. 7. Comparison with AMCP. (a) Throughput versus number of channels. (b) Throughput versus traffic load. Four channels.

7 .CONCLUSION

Here, we have introduced DISH, which is a distributed flavor of control-plane cooperation, as a new approach to wireless protocol design. It enables transmitter-receiver pairs to exploit the knowledge at individual idle neighbors to make more informed decisions in communication. Applying DISH to multichannel ad hoc networks, we propose a cooperative multichannel MAC protocol called CAM-MAC, where idle neighbors share control information with transmitter-receiver pairs to overcome MCC problems. This protocol uses a single transceiver and, unlike many other protocols, is fully asynchronous.

The simple idea of DISH turns out to be very effective. In the comparison of CAM-MAC with and without DISH, we observe remarkable performance difference. In the comparison with three recent and representative multichannel MAC protocols, MMAC, SSCH, and AMCP, CAM-MAC significantly outperforms all of them. Our implementation on COTS hardware and experiments further validates the advantages of CAM-MAC and the DISH idea.

In a sense, DISH enables nodes to store channel usage information at their neighbors, and retrieve this information when it is needed. We also highlight that this is not a compulsory coordination mechanism, a network does not rely on cooperation and still operates when cooperation is not available. Ultimately, we believe that control-plane cooperation merits due consideration in the future design of wireless network protocols.

8. ACKNOWLEDGEMENT

We take immense pleasure in thanking our chairman Dr. Jeppiaar M.A, B.L, Ph.D, the Directors of Jeppiaar Engineering College Mr. Marie Wilson, B.Tech, MBA, (Ph.D), Mrs. Regeena Wilson, B.Tech, MBA, (Ph.D)and the principal Dr. Sushil Lal Das M.Sc(Engg.), Ph.D for their continual support and guidance. We would like to extend our thanks to my guide, our friends and family members without whose inspiration and support our efforts would not have come to true. Above all, we would like to thank God for making all our efforts success.

9. REFERENCES

[1] T. Luo, M. Motani, and V. Srinivasan, “CAM-MAC: A Coopera-tive Asynchronous Multi-Channel MAC Protocol for Ad Hoc Networks,” Proc. IEEE Third Int’l Conf. Broadband Comm., Networks and Systems (BROADNETS ’06), Oct. 2006.

[2] J. So and N. Vaidya, “Multi-Channel MAC

for Ad Hoc Networks: Handling Multi-Channel Hidden Terminals Using a Single Transceiver," Proc. ACM MobiHoc, 2004.

[3] P. Bahl, R. Chandra, and J. Dunagan, "SSCH: Slotted Seeded Channel Hopping for Capacity Improvement in IEEE 802.11 Ad-Hoc Wireless Networks," Proc. ACM MobiCom, 2004.

[4] J. Shi, T. Salonidis, and E.W. Knightly, "Starvation Mitigation through Multi-Channel Coordination in CSMA Multi-Hop Wire-less Networks," Proc. ACM MobiHoc, 2006.

[5] S.-L. Wu, C.-Y. Lin, Y.-C. Tseng, and J.-P. Sheu, "A New Multi-Channel MAC Protocol with On-Demand Channel Assignment for Multi-Hop Mobile Ad Hoc Networks," Proc. Int'l Symp. Parallel Architectures, Algorithms and Networks (ISPAN), 2000.

[6] A. Nasipuri, J. Zhuang, and S.R. Das, "A Multichannel CSMA MAC Protocol for Multihop Wireless Networks," Proc. IEEE Wireless Comm. and Networking Conf. (WCNC), 1999.

[7] A. Nasipuri and J. Mondhe, "Multichannel CSMA with Signal Power-Based Channel Selection for Multihop Wireless Networks," Proc. IEEE Vehicular Technology Conf. (VTC), 2000.

[8] N. Jain, S.R. Das, and A. Nasipuri, "A Multichannel CSMA MAC Protocol with Receiver-Based Channel Selection for Multihop Wireless Networks," Proc. 10th Int'l Conf. Computer Comm. and Networks (ICCCN), 2001.

[9] A. Adya, P. Bahl, J. Padhye, and A. Wolman, "A Multi-Radio Unification Protocol for IEEE 802.11 Wireless Networks," Proc. IEEE First Int'l Conf. Broadband Comm., Networks and Systems (BROADNETS), 2004.

AUTHORS PROFILE

P. Jesu Jayarin B.E., M.E., (Ph.D) working as an Assistant Professor at Jeppiaar Engineering College, Chennai and he has more than 5 years of teaching experience. His areas of specializations are Mobile Computing, Computer Networks, Network security and TCP/IP.

Dr. T. Ravi, B.E, M.E, Ph.D is a Professor & Head of the Department of CSE at KCG college of Technology, Chennai. He has more than 18 years of teaching experience in various engineering institutions .He has published more than 20 papers in International Conferences & Journals.

KEY GENERATION FOR AES USING BIO-METIC FINGER PRINT FOR NETWORK DATA SECURITY

1. Dr.R.Seshadri
Director
University Computer Center
Sri Venkateswara University, Tirupati
ravalaseshadri@gmail.com

2. T.Raghu Trivedi
Research Scholar,
Department of Computer Science
Sri Venkateswara University, Tirupati.
tamirisa_t1@yahoo.com

Abstract

Encryption is one of the Essential security technologies for computer data, and it will go a long way toward securing information. The unauthorized thefts in our society have made the requirement for reliable information security mechanisms. Even though information security can be accomplished with the help of a prevailing tool like cryptography, protecting the confidentiality of the cryptographic keys is one of the significant issues to be deal with. Here we proposed a biometric-crypto system which generates a cryptographic key from the Finger prints for encrypting and decrypting the information the popular biometric used to authenticate a person is fingerprint which is unique and permanent through out a person's life. Hence, the biometric is gone eternally and possibly for all the applications. If your information traverses on net to reach destination a number of attacks may be done. To protect from attacks we proposed a system which will encrypt the data using AES with biometric based key generation technique.

Key Words: *Decryption, Encryption, Histogram Equalization, Minutiae points, Morphological Operation.*

1. Introduction

Here we are using the AES for the encryption and Decryption process. For AES Key is important. Protecting the confidentiality of the cryptographic keys is one of the significant issues. We generate the key from the receiver's finger print template. The security to that is provided with the help of finger print of the sender.

Biometric Cryptographic Key Generators, or BKGs, follow a similar design: during an enrollment phase, biometric samples from a user are collected; statistical functions, or *features*, are applied to the samples; and some representation of the output of these features is stored in a data structure called a biometric *template*. Later, the same user can present another sample, which is processed with the stored template to reproduce a key [1].

Initially, the fingerprints are employed to extract the minutiae points which are transformed in an efficient manner to obtain deformed points. Subsequently, the deformed points are employed to generate the cancelable templates which are utilized for the extraction of irrevocable keys.

One way to protect privacy is to encrypt the information we used a system which will encrypt the data using AES with Novel method of biometrics based key generation technique. Biometric crypto systems can operate in one of the following three modes 1.Key Release 2.Key binding 3.Key generation.

In the key Release mode Biometric authentication is decoupled from the key release mechanism. The biometric template and key are stored as separate entities and the key is released only if the biometric matching is successful

.In the key binding mode the key and the templates are monolithically bound with in a cryptographic frame work. It is computationally infeasible to decode the key or template without any knowledge of the user's biometric data. A crypto biometric matching algorithm is used to perform authentication and key release in a single step.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

In the key generation mode the key is derived directly from the biometric data and is not stored in the data base.

Though it is easy to implement a biometric crypto system in the key release mode such a system is not appropriate for a high security application because the biometric template is not secure template security is critical issue because stolen templates cannot be revoked. Key binding mode are more secure but difficult to implement due to large intra class variations in biometric data i.e. samples of the same biometric trait of user obtained over a period of time differ substantially.

2.AES

Encryption is one of the securities Technique for the information/data traversing through network. Encryption is process of transforming information using an algorithm to make it unreadable (cipher) to any one except sender and receiver. An encryption algorithm along with a key used in encryption and decryption of data .AES is one of the popular Algorithms used in symmetric key cryptography [3, 6]. It is symmetric block cipher that can encrypt/decrypt information. It is used at top secret level.AES supports key sizes 128,192 and 256 bits and will serve as replacement for the DES, which has a less key size. AES can encrypt data much faster than DES enhancement (Triple DES)..

- **Cipher text=E_K(plain text)**
- **Plain text= D_K(cipher text)**
- **$D_K(E_K$(plain text))=plain text**

3. Cryptography and biometrics

Cryptography provides security to the information which is transferring over the insecure channel. Here key will play a major role, because it is used for encryption as well as for decryption. Lengthy key used to encrypt and decrypt sending and receiving the messages respectively. These keys can be guessed/ cracked. Maintaing and sharing the lengthy keys is critical problem. This can be overcome with the help of biometric system [5].

There are different biometric techniques. We concentrate on fingerer prints to generate key.

Key vector Is formed based on minutiae points (ridge ending and ridge bifurcation)

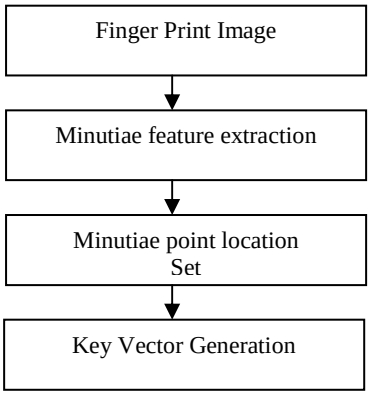


Figure 1: Generating key vector from finger print

4. Key Generation from Finger Print

Here we have to extract the minutiae points from the fingerprint to generate the key. The process is as below

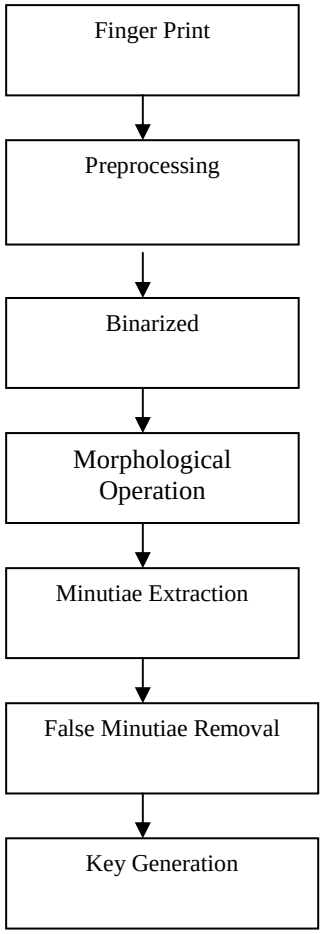


Figure2: Key generation from minutiae points

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

4.1 Extraction of Minutiae Points from Fingerprints

The extraction of minutiae points from the fingerprint image is discussed in this section. It is supposed that fingerprints are distinct across individuals and across the fingers of a particular individual [1] . Since many existing fingerprint authentication systems are based on minutiae points, which are feature points extracted from a raw fingerprint image, we have employed the minutiae points in our scheme as well.

A fingerprint can be defined as a pattern of ridges and valleys on the tip of the finger. A fingerprint is therefore described by the distinctiveness of the local ridge features and their relationships. Minutiae points denote these local ridge characteristics that appear either at a ridge ending or a ridge bifurcation. The point where the ridge comes to an abrupt end is known as ridge ending and the ridge bifurcation is denoted as the point where the ridge divides into two or more branches [1]. The major steps involved in the minutiae points' extraction are as follows:

• Segmentation
• Minutiae Extraction

4.1.1Segmentation

The first step in the minutiae points' extraction is segmentation. The input fingerprint image is segmented from the background to actually extract the region comprising the fingerprint, which ensures the removal of noise. Segmentation of an image represents the division or separation of the image into regions that have similar attributes. At first, the image is pre-processed. The pre-processing phase includes the following: histogram equalization. Later, the pre-processed image is divided into blocks and segmentation is carried out. The sample fingerprint images are shown

Figure 3: Two Sample Fingerprint Images

4.1.1.1 Pre-Processing

The pre-processing of fingerprint images includes Histogram Equalization

4.1.1.1. a. Histogram Equalization

Histogram equalization amplifies the local contrast of the images, particularly when they are represented with very close contrast values. It is possible to distribute intensity through the histogram with the aid of this regulation. Histogram equalization utilizes a monotonic, non-linear mapping that re-assigns the intensity values of pixels in the input image in such a manner that the output image comprises a uniform distribution of intensities (i.e. a flat histogram). The original histogram of a Fingerprint image is of bimodal type, the histogram after the histogram equalization transforms all the range from 0 to 255 which results if enhanced visualization effect [7]. The results of histogram equalization are depicted in Figure 4.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Figure 4: Fingerprint Images after Histogram Equalization

4.1.2 Minutiae Points Extraction

Finally, the minutiae points are extracted from the enhanced fingerprint image. The steps involved in the extraction of minutiae points are as follows:

- Binarization
- Morphological Operations
- Minutiae points' extraction

Initially, the enhanced image is binarized. After binarization, morphological operations are performed on the image to remove the obstacles and noise from it. Finally, the minutiae points are extracted using the approach discussed.

4.1.2.1 Binarization

The binary images with only two levels of interest: The black pixels that denote ridges and the white pixels that denote valleys are employed by almost all minutiae extraction algorithms. A grey level image is translated into a binary image in the process of binarization, by which the contrast between the ridges and valleys in a fingerprint image is improved. Hence, the extraction of minutiae is achievable. The grey-level value of every pixel in the enhanced image is analyzed in the binarization process. Then, the pixel value is set to a binary value one when the value is greater than the global threshold, or else a zero is set as the pixel value. The foreground ridges and the background valleys are the two level of information held by the ensuing binary image. Removal of distortions present in the image is performed followed by the retrieval of the exact skeleton image from the image.

4.1.2.2 Morphological Operation

The binary morphological operators are applied on the binarized fingerprint image. Elimination of any obstacles and noise from the image is the primary function of the morphological operators. Furthermore, the unnecessary spurs, bridges and line breaks are removed by these operators. Then thinning process is performed to reduce the thickness of the lines so that the lines are only represented except the other regions of the image. Clean operator, Hbreak operator, Spur operator and Thinning are the morphological operators applied.

4.1.2.3 Minutiae points' extraction

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

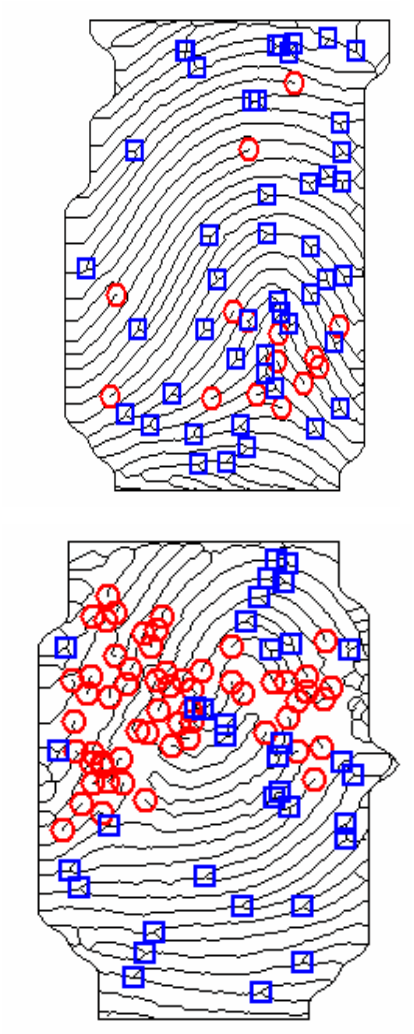


Figure 5: Fingerprint Images with Minutiae Points

Ridge Thinning is to eliminate the redundant pixels of ridges till the ridges are just one pixel wide. uses an iterative, parallel thinning algorithm. In each scan of the full fingerprint image, the algorithm marks down redundant pixels in each small image window (3x3). And finally removes all those marked pixels after several scans. After the fingerprint ridge thinning, marking minutia point is relatively easy .For each 3X3 window, if the central pixel is 1 and has exactly 3 one-value neighbors, then the central pixel is a ridge branch. if the central pixel is 1 and has only 1 one –value neighbors , then the central pixel is ridge ending. Suppose both the uppermost pixel with value 1 and the rightmost pixel with value 1 have another neighbor outside the 3X3 window, so the two pixels will be marked as branches too. But actually only one branch is located in the small region. So a check routine requiring that none of the neighbors of a branch are branches is added.

4.1.2.4 False Minutia Removal:

The preprocessing stage does not totally heal the fingerprint image .For example, false ridge breaks due to insufficient amount of ink and ridge cross-connections due to over inking are not totally eliminated. Actually all the earlier stages themselves occasionally introduce some artifacts which later lead to spurious minutia. These false minutias will significantly affect the accuracy of matching if they are simply regarded as genuine minutia. So some mechanisms of removing false minutia are essential to keep the fingerprint verification system effective.

Key Generation From Minutiae Points

In this section we explain the **key Generation Algorithm** [2] **Assumptions**

Kl	Length of the AES key
Mp	Minutiae point set
Kl	key length
Np	Size of Minutiae pint set
S	Seed value
Sl	seed limit
M	(x, y)-co-ordinate of a minutiae point
Kv	Key Vector

Step 1: The Extracted minutiae pints are represented

As

$Mp=\{mi\}\ i=1\ldots..\ Np$

Step2: The initial key vector is defined as follows

$Kv = \{xi: p(xi)\}\ i=1\ldots\ldots\ldots Kl$

Where $p(x) = Mp[I\%\ Np] + Mp[(i+1)\ \%\ Np] + S$

$i=1\ldots\ldots\ldots..Kl$

step3: Initial value of S is equal to total Number of Minutiae pints . The value of S will be dynamically changed as follows

$S=Kv(i)\ \%Sl, -1<i<Kl$

Step4: Initial key vector (Kv) is converted in to a

Matrix Km of size Kl / 2 * Kl / 2

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Km= (aij) Kl / 2 * Kl / 2

Step 5:

A intermediate key vector is generated as follows KIV= {Ki: (m (ki)} i=1…..Kl

Where

M (k) =|Aij|,

Aij=Km i, j: i + size + j + size, -1<i<kl/2

Aij is a sub matrix formed from the key matrix

Step 6: Final key vector is formed is

$$Sc = \begin{cases} 1, \text{ if KlV [i]> mean (KlV)} \\ 0, \text{ other wise} \end{cases}$$

5. Conclusion

From the above discussion I have proposed a method for providing security to the information transferring on the network using Encryption and a novel approach for fingerprint based cryptography system. Here we used fingerprint patterns, which are stable through out person's lifetime. Since it creates more complexity to crack or guess the crypto keys. This approach has reduced the complicated sequence of the operation to generate crypto keys as in the traditional cryptography system.

6. References

1. N.Lalithamani, K.P.Soman "Irrevocable Cryptographic Key Generation from Cancelable Fingerprint Templates: An Enhanced and Effective Scheme". European Journal of Scientific Research ISSN 1450-216X Vol.31 No.3 (2009), pp.372-387
2. P.Arul, Dr.A.Shanmugam "Generate a Key for AES Using Biometric For VOIP Network Security" Journal of Theoretical and Applied Information Technology 2009.107-112.
3. "Advanced Encryption Standard "from http://en.wikipedia.org/wiki/Advanced_Encryption_ Standard
4. Jain, A.K., Ross, A. and Prabhakar, S, "An introduction to biometric recognition," IEEE Transactions on Circuits and Systems for Video Technology, Vol. 14, No. 1, pp: 4- 20, 2004.
5. Umut Uludag, Sharath Pankanti, Salil Prabhakar, Anil K .Jain "Biometric Cryptosystems Issues and Challenges" Proceedings of the IEEE 2004.
6. Announcing the "ADVANCED ENCRYPTION STANDARD (AES)" –Federal Information, Processing Standards Publication 197, November 26, 2001
7. Jain, A.K.; Prabhakar, S.; Hong, L.; Pankanti, S., "Filter bank-based fingerprint matching", IEEE Transactions on Image Processing, vol. 9, no. 5, pp: 846-859, May 2000, Doi:10.1109/83.841531.

Authors Profile

Dr.R.Seshadri was born in Andhra Pradesh, India, in 1959. He received his B.Tech degree from Nagarjuna University in 1981. He completed his Masters degree in Control System Engineering from PSG College of Technology, Coimbatore in 1984. He was awarded with PhD from Sri Venkateswara University, Tirupati in 1998. He is currently Director, Computer Center, S.V.University,Tirupati, India. He published number of papers in national and international journals. At present 12 members are doing research work under his guidance in different areas.

Mr.T.RaghuTrivedi received MCA degree from Andhra University, Vizag He received his M.Tech in Computer Science from Nagarjuna University.He is a research scholor in S.V.University Tirupati, Andhra Pradesh.

Classification of Five Mental Tasks Based on Two Methods of Neural Network

Vijay Khare[1]
Jaypee Institute of Information Technology
Dept. of Electronics and Communication, Engineering
Nioda, India
Email : vijay.khare@jiit.ac.in

Jayashree Santhosh[2]
Indian Institute of Technology,
Computer Services Centre
Delhi, India
Email : jayashree@cc.iitd.ac.in

Sneh Anand[3]
Indian Institute of Technology,
Centre for Biomedical Engineering Centre
Delhi, India
Email : sneh@iitd.ernet.in

Manvir Bhatia[4]
Sir Ganga Ram Hospital,
Department of Sleep Medicine,
New Delhi, India
Email : manvirbhatia1@yahoo.com

Abstract— **In this paper performance of two classifiers based on Neural Network were investigated for classification of five mental tasks from raw Electroencephalograph (EEG) signal. Aim of this research was to improve brain computer interface (BCI) system applications. For this study, Wavelet packet transform (WPT) was used for feature extraction of the relevant frequency bands from raw electroencephalogram (EEG) signals. The two classifiers used were Radial Basis Function Neural Network (RBFNN) and Multilayer Perceptron Back propagation Neural Network(MLP-BP NN) . In MLP-BP NN five training methods used were (a) Gradient Descent Back Propagation (b) Levenberg-Marquardt (c) Resilient Back Propagation (d) Conjugate Learning Gradient Back Propagation and (e) Gradient Descent Back Propagation with movementum.**

Index Terms— **Electroencephalogram (EEG), Wavelet Packet Transform (WPT), Radial Basis Function Neural Network(RBFNN), Multilayer Perceptron back propagation Neural Network(MLP-BP NN), Brain computer interfaces (BCI).**

I. INTRODUCTION

Brain signals extracted through EEG carry information needed for the design and development of brain computer interface (BCI) systems. It is well documented that proper feature extraction and classification methods are the key features deciding the accuracy and speed of BCI systems [1-5]. ANN has been more widely accepted as one of the best classification method to distinguish various mental states from relevant EEG signals.

Past two decades have witnessed the importance of innovative BCI with voice, vision and a combination of these, as a communication platform [6-9]. Effective attempts have been made to achieve successful BCI systems based on bioelectric signals. They were mainly to help patients with various neuromuscular disorders by providing them a way of communication to the world, through extracting information from their intensions. So far the accuracy of the classification has been one of the main pitfalls of the existing BCI systems, since it directly affects the decision made as the BCI output. The speed &accuracy could be improved by implementing better methods for feature extraction and classification [10-16]. In this study, wavelet packet transform (WPT) method was used to capture the information of mental tasks from eight channel EEG signals of nine subjects. The coefficients of wavelet packet transform (WPT) were used as the best fitting input vector for clssifiers [17]. The two classifiers (RBFNN and MLP-BP NN) were used to compare the performance to discriminate five mental tasks**.**

II. METHODOLOGY

A. Subjects

Nine right-handed healthy male subjects of age (mean: 23yr) having no sign of any motor- neuron diseases were selected for the study. A pro-forma was filled in with detail of their age & education level as shown in table1. The participants were student volunteers for their availability and interest in the study. EEG data was collected after taking written consent for participation. Full explanation of the experiment was provided to each of the participants.

TABLE 1: CLINICAL CHARACTERISTICS OF SUBJECTS

S.No.	Subject	Age	Educational status
1	Subject 1	22	BE
2	Subject 2	21	BE
3	Subject 3	23	BE
4	Subject 4	27	M.TECH
5	Subject 5	23	BE
6	Subject 6	22	BE
7	Subject 7	27	M.TECH
8	Subject 8	22	BE
9	Subject 9	22	BE

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

B. EEG Data Acquisition

EEG Data used in this study was recorded on a Grass Telefactor EEG Twin3 Machine available at Deptt. of Neurology , Sir Ganga Ram Hospital, New Delhi. EEG recording for nine selected subjects were done for five mental tasks for five days. Data was recorded for 10 sec during each task and each task was repeated five times per session per day. Bipolar and Referential EEG was recorded using eight standard positions C3, C4, P3, P4, O1 O2, and F3, F4 by placing gold electrodes on scalp, as per the international standard 10-20 system of electrode placement as shown in figure 1. The reference electrodes were placed on ear lobes and ground electrode on forehead. EOG (Electooculargram) being a noise artifact, was derived from two electrodes placed on outer canthus of left and right eye in order to detect and eliminate eye movement artifact. The settings used for data collection were: low pass filter 1Hz, high pass filter 35 Hz, sensitivity 150 micro volts/mm and sampling frequency fixed at 400 Hz.

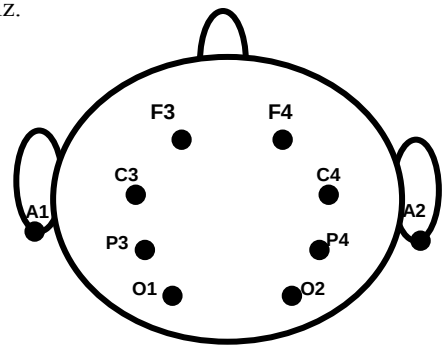


Figure1:- Montage for present study

C. Experiment Paradigm

An experiment paradigm was designed for the study and the protocol was explained to each participant before conducting the experiment. In this, the subject was asked to comfortably lie down in a relaxed position with eyes closed. After assuring the normal relaxed state by checking the status of alpha waves, the EEG was recorded for 50 sec, collecting five session of 10sec epoch each for the relaxed state. This was used as the baseline reference for further analysis of mental task. The subject was asked to perform a mental task on presentation of an audio cue. Five session of 10sec epoch for each mental task were recorded, each with a time gap of 5 minute (as shown in figure2). The whole experiment lasted for about one hour including electrode placement.

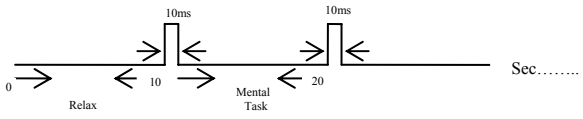


Figure 2: Timing of the Protocol

Data collected from nine subjects performing five mental tasks were analyzed. The following mental tasks were used to record the appropriate EEG data.

- Movement Imagination:-The subject was asked to plan movement of the right hand.
- Geometric Figure Rotation:-The subject was given 30 seconds to see a complex three dimensional object, after which the object was removed. The subject was instructed to visualize the object being rotated about an axis.
- Arithmetic Task:-The subject was asked to perform trivial and nontrivial multiplication. An example of a trivial calculation is to multiply 2 by 3 and nontrivial task is to multiply 49 by 78. The subject was instructed not to vocalize or make movements while solving the problem.
- Relaxed: - The subject was asked to relax with eyes closed. No mental or physical task to be performed at this stage.

D. Feature Extraction

The frequency spectrum of the signal was first analyzed through Fast Fourier Transform (FFT) method. The FFT plot of signals from the most relevant electrode pairs were observed along with average change in EEG power for each mental tasks as shown in figure (2-6).

For relaxed, the peaks of power spectrum almost coincide (or difference of (0-10 %) for central and occipital area in the alpha frequency range (8-13Hz). EEG recorded with relaxed state is considered to be the base line for the subsequent analysis. Mu rhythms are generated over sensorimotor cortex during planning a movement. For movement imagery of right hand, maximum upto 50% band power attenuation was observed in contralateral (C3 w.r.t C4) hemisphere in the alpha frequency range (8-13Hz). For geometrical figure rotation, the peak of the power spectrum was increased in right hemisphere rather than left in the occipital area for the alpha frequency range (8-13Hz). For both trivial and nontrivial multiplication, the peak of the power spectrum was increased in left hemisphere rather than right hemisphere in the frontal area for the alpha frequency range (8-13Hz).

The data was preprocessed using Wavelet packet transform to extract the most relevant information from the EEG signal. [18-19]. By applying Wavelet packet transform on the original signal wavelet coefficients in the (8-13Hz) frequency band at the 5^{th} level node (5, 3) were obtained. We were able to reduce 1 second of EEG data to 21 coefficients. The signal was reconstructed at node (5, 3). These coefficients are scaled and used as the best fitting input vector for classifiers.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

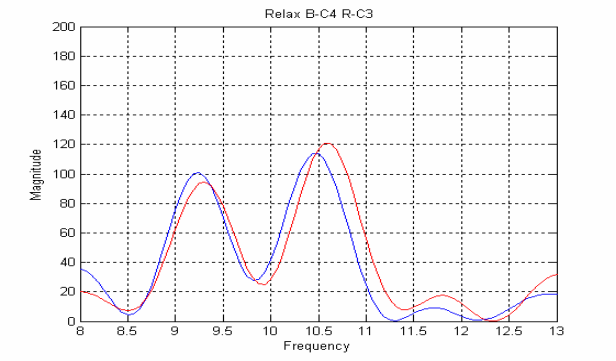


Fig 2:- Power Spectra for Relax state at C3 and C4 channel

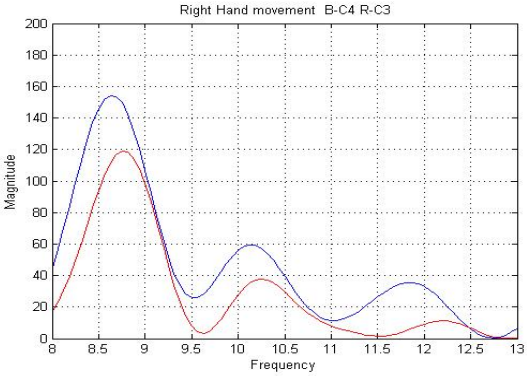


Fig 3:- Power Spectra for planning of right hand movement at C3 and C4

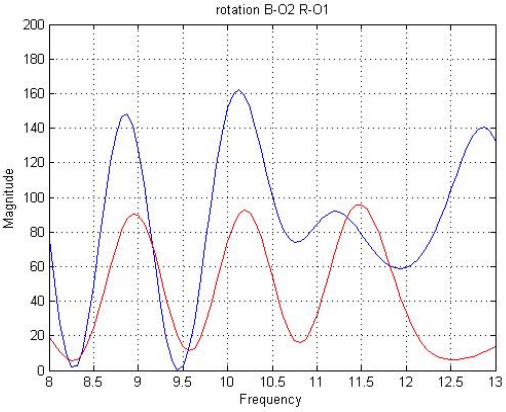


Fig 4:- Power Spectra for visual rotation at O1 and O2 channel

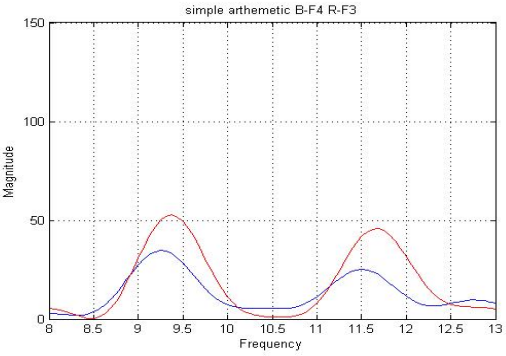


Fig 5:- Power Spectra for Simple Arithmetic F3 and F4 channel

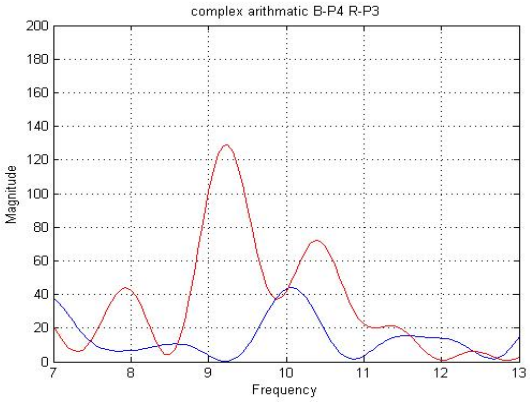


Fig 6:- Power Spectra for complex Arithmetic P3 and P4 channel

E. Classifier

The main advantage of choosing artificial neural network for classification was due to fact that ANN's could be used to solve problems, where description for the data is not computable. ANN could be trained using data to discriminate the feature. We have compared two classifiers Radial Basis Function Neural Network (RBFNN) and Multilayer Perceptron Back propagation Neural Network (MLP-BP NN). In MLP-BP NN five training methods used were (a) Gradient Descent Back Propagation (b) Levenberg-Marquardt (c) Resilient Back Propagation (d) Conjugate Learning Gradient Back Propagation and (e) Gradient Descent Back Propagation with movementum.

- Radial Basis Function Neural Network

For this (RBFNN) classifier, a two layer network was implemented with 21 input vectors, a hidden layer consisting as many as hidden neurons as there are input vectors with Gaussian activation function and one neuron in the output layer [20-23]. RBFNN produce a network with zero error on training vectors. The output neuron gives 1 for a mental task and 0 for relax task.

- Multilayer Perceptron Back Propagation Neural Network

For this classifier, a two layer feed forward neural network was used with topology of {10, 1}. 10 neurons in hidden layer and 1 neuron in output layer. The neural network was designed to accept a 21 element input vector and give a single output. The output neuron was designed to give 0 for baseline (relax task) and 1 for mental task. The five different training methods used for this classifier were Gradient Descent, Resilient Back propagation, Levenberg-Marquardt, Conjugate Gradient Descent and Gradient Descent back propagation with movementum [24-27]. Parameter used for five training methods of neural network for classification of five mental tasks as shown in the table 2.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

TABLE 2
PARAMETER USED FOR DIFFERENT ALGORITHMS
WITH TOPOLOGY { 10, 1}

Gradient descent with momentum (GDM)	
Topology {10,1}	A=.01. Mu = 0.01
MSE= 1e-5	
Epoach=5000	
Gradient descent method(GDBP)	
Topology {10,1}	A=.01
MSE=1exp-(5)	
Epoach=5000	
Resilient Back propagation(RBP)	
Topology {10,1}	A=.01
MSE=1exp-(5)	
Epoach=5000	B=0.75 and β_1=1.05
Conjugate gradient descent(CGBP)	
Topology {10,1}	A=.01
MSE=1exp-(5)	
Epoach=5000	
Levenberg-Marquardt(LM)	
Topology {10,1}	Mu=.01
MSE=1exp-(5)	
Epoach=5000	Mu_dec=0.1 and Mu_inc =10

F. Performance

The study evaluated the performance of two classifiers (RBFNN and MLP-BP NN) for classification of five mental tasks. For MLP-BP NN classifier five different training methods used were Gradient Descent, Resilient Back propagation, Levenberg-Marquardt, Conjugate Gradient Descent and Gradient Descent back propagation with movementum. 60% of entire EEG data (five sessions, five mental tasks with nine subject) was taken as training data. Remaining 40% of EEG data was taken as test data and the performances were recorded. The entire analysis of the recorded data was carried out using Matlab® 7.0 from Mathworks Inc., USA.

Performance (R_C) is calculated in percentage (%) as ratio between correctly classified patterns in the test set to the total number of patterns in the test set [28].

$$Rc = \frac{\text{Number of correctly classified test patterns}}{\text{Total number of patterns in the test set}}$$

III. RESULT AND DISCUSSION

Nine right-handed male subjects participated in the experiments. The subjects were asked to perform five mental tasks namely relaxed, movement imagery, geometrical figure rotation and arithmetic task (trivial and non trivial multiplication). Out of 50 sec data recorded data the most relevant one second epoch of signal were used for classification of each mental task. WPT is an excellent signal analysis tool, especially for non stationary signals. Hence in the present study, WPT was used for feature extraction [29].

As per literature most prominent area in brain for domain of information during five mental tasks was shown in table3. Amplitude of power spectrum almost coincides in central and occipital area at a particular base frequency (8-13Hz) for relaxed states [12].

The frequency spectrum of the signal was observed that the amplitude of the power spectrum for alpha frequency range (8-13Hz) had an attenuation in contralateral area for movement imagery task [12 17].

For geometrical figure rotation, It was observed that the amplitude of the power spectrum increased in the right occipital region for alpha frequency range (8-13Hz) [30 31].

For trivial multiplication, it was observed that the amplitude of the power spectrum increased in the left frontal region for alpha frequency range (8-13Hz).For nontrivial multiplication, it was observed that the amplitude of the power spectrum increased in the left parietal region for alpha frequency range (8-13Hz) [31].

For MLP-BP NN classifier, Resilient back propagation training method showed better performance than other back propagation training methods (Gradient Descent method Levenberg-Marquardt, Conjugate Gradient Descent and Gradient Descent back propagation with movementum) for classification of five mental tasks (As shown in table 4).

TABLE 3: DOMAIN OF INFORMATION

Tasks	**Domain of information**	**(Contralateral/ Ipsilateral)**	**Type of change in amplitude of alpha rhythm(8-13Hz)**
Base line	Occipital, Central	Contralateral	Coincide
Movement Imagination	Central	Contralateral	Decreased
Geometrical figure rotational	Occipital	Ipsilateral	Increased
Arithmetic's(trivial and non trivial) operation	Frontal, parietal	Ipsilateral	Increased

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

The present study was a comparison of two classifier, MLP-BP NN with Resilient back propagation training method and RBFNN to discriminate five mental tasks effectively. From figure(7-12) we can say that RBF Neural Network has better performance as compare to MLP-BP NN with Resilient back propagation method for classification of five mental tasks movement imagination(M), figure rotation(R), simple arithmetic (SA) task and complex arithmetic (CA) task, w.r.t baseline. Average accuracy was obtained 100% by using RBFNN classifier.

IV. CONCLUSION

For various applications of BCI systems, it is necessary that EEG feature related to the human intentions were to be uniquely identified as accurate as possible. In this paper, the two classifiers used were, MLP-BP NN with resilient back propagation training method and RBFNN**.** Radial basis function neural network was found to be most suitable for classification of five mental tasks.

Radial basis function networks showed a better capability for solving larger data size problems at fast learning speed, because of their capability of local specialization and global generalization. Various other classification methods could be analyzed in future for better performance.

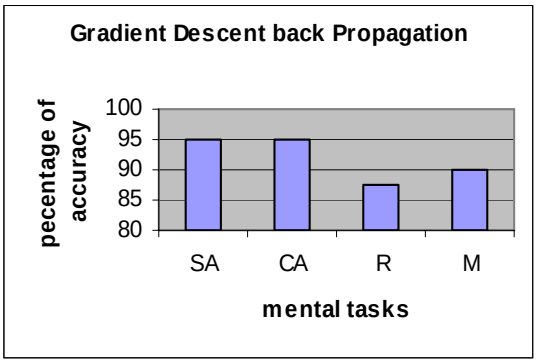


Fig 7: Classification accuracy using GD BP training methods

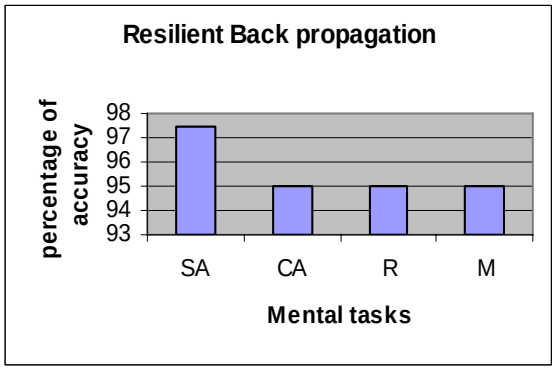


Fig 8: Classification accuracy RBP training methods

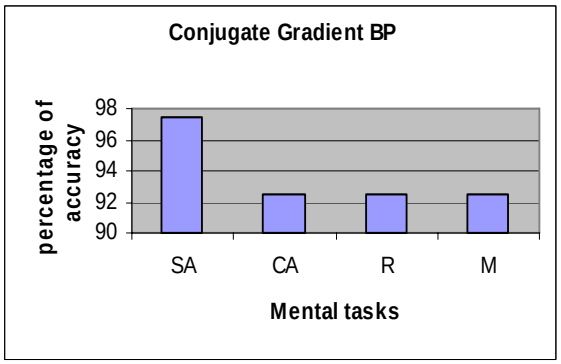


Fig 9: Classification accuracy CGBP training methods

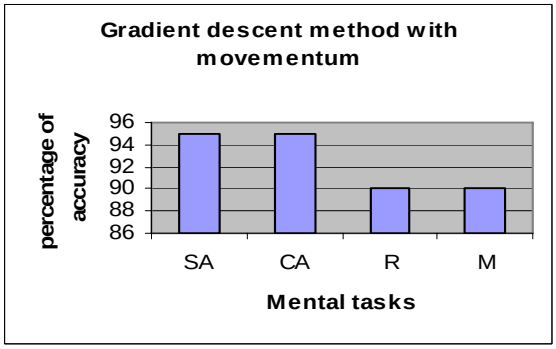


Fig 10: Classification accuracy GDM training methods

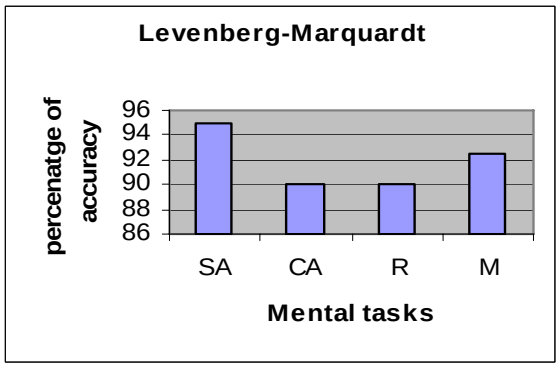


Fig 11: Classification accuracy LM training methods

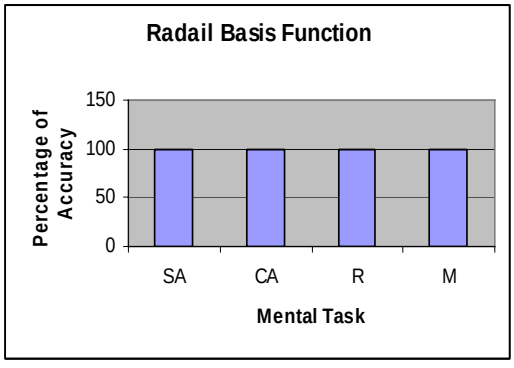


Figure 12: Classification accuracy for RBFNN

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

ACKNOWLEDGMENT

The authors would like to acknowledge their gratitude to the scientific and technical staff of EEG Laboratory of Sir Ganga Ram hospital, New Delhi for the help in carrying out the experiment.

REFERENCES

[1] Lotte F., M. Congedo, A. Lecuyer, F. Lamarche, B Arnaldi A Review of Classification algorithms for EEG bases brain computer interface Journal of neural Engineering, vol. 4, R1-R13, 2007.

[2] .Wolpaw J.R , N.Birbaumer, D.J Mc Farland, G.Plurtscheller, T.M.Vaughan, Brain computer Interfaces for communication and control Clinical Nurophys. , 113,767-791, 2002.

[3] Pfurtschelle G. , D.Flotzinger, and J.Kalcher Brain Computer interface-A new communication device for handicapped people J Microcomput. Applicate., vol. 16,. pp.293- 299,1993.

[4] Wolpaw J.R, T.M.Vaughan and E.Donchin EEG Based Communication prospects and problems IEEE,Trans.Rehab.Eng.vol.4,.pp.425-430, Dec.1996.

[5] Keirn Z. A. and J. I. Aunon, A new mode of communication between man and his surroundings IEEE Trans. Biomed. Eng., vol. 37, no. 12, pp. 1209–1214, Dec. 1990.

[6] Wolpaw J.R., G.E. Leob, B.Z. Allison, E.Donchin, J. N. Turner BCI Meeting 2005-Wokshop on Signals and Rerecording Methods IEEE transactions on Neural systems and rehabilitation Engineering, vol.14, no 2, June 2006.

[7] Anderson Charles W., Erik A. Stolz, Sanyogita Shamsunder Multivariable Autoregressive Model for classification of Spontaneous Electroencephalogram During Mental Tasks IEEE Trans. Biomed. Eng., vol. 45 no-3, pp.277-278, 1998.

[8] Elean A, Curran and Jamaica Strokes learning to control brain activity A Review of the production and control of EEG components for driving Bain computer interface systems Brain and cognition, 51,326-336, 2003.

[9] Wolpaw J. R., D. J. Mc Farland, T.M. Vaughan The wads worth Centre Brain Computer interface research and Development Program IEEE Trans. on Neural System and rehab. Eng., 11(2)) 204-207, 2003.

[10] Liu Hailong, Jue Wang, C.Zheng and Ping He Study on the Effect of Different Frequency band of EEG Signal on the Mental Tasks Classification IEEE Engineering in medicine and biology 27th annual conference shanghai china, sept.1-4, and 2005.

[11] Boostani R. , B. Graimann, M.H.Moradi, G. Plurfscheller Comparison approach toward finding the best feature and classifier in cue BCI Med. Bio. Engg., Computer 45,403-413, 2007.

[12] Pfurtscheller G., C. Neuper, A. Schlogl, and K. Lugger Separability of EEG signals recorded during right and left motor imagery using adaptive auto regressive parameters IEEE. Trans. on rehabilitation Engineering, 6(3), pp.316-325, 1998.

[13] Palaniappan R. Brain computer interface design using band powers extracted during mental task proceeding of the 2ndInternational IEEE EMBS Conference on Neural Engineering 321- 324, 2005.

[14] Kouhyar Tavakolian, Faratash Vasefi, Siamak Rezaei, Mental Task Classification for Brain Computer Interface Application, first Canadian student conference on biomedical 2006.

[15] Palaniappan Ramaswamy Utilizing Gamma Band to Improve Mental task based Brain –Computer Interface Design IEEE Trans. on Neural systems and rehabilitation Engg. , vol.14, no 3, September 2006.

[16] Lee J.C, D.S Tan Using a low cost Electroencephalograph for the Mental task classification in HCI Research UIST 06, October 15-18, montreux, Switzerland 2006.

[17] Santhosh Jayashree, Manvir Bhatia, S.Sahu, Sneh Anand Quantitative EEG analysis for assessment to plan a task in ALS patients, a study of executive function (planning) in ALS, Cognitive brain research 22, 59-66, 2004.

[18] Akay M Wavelet in biomedical engineering, annalsin Biomedical Engineering 23 (5), 531-542, 1995.

[19] Strang Gilbert and Troung Ngugen. Wavelet and Filter Banks Wellesley Cambridge press, 1997.

[20] Larsson Elisabeth, Krister Åhlander, and Andreas Hall Multi-dimensional option pricing using radial basis functions and the generalized Fourier transform, In J. Comput. Appl. Math., 2008.

[21] Pettersson Ulrika, Elisabeth Larsson, Gunnar Marcusson, and Jonas Persson Improved radial basis function methods for multi-dimensional option pricing, In J. Comput. Appl. Math., 2008.

[22] Warwick K. , J. Mason ., and Sutanto E Centre Selection for Radial Basis Function Network", Artificial Network and Genetic algorithms proceeding of the international conference ,Ales, France ,D. Pearson,N.Seetle and R. Albrecht(Eds) 309-312, 1995.

[23] Chen S. , C.F.N. Cowan, and P. M. Grant Orthogonal Least Squares Learning Algorithm for Radial Basis Function Networks IEEE Transactions on Neural Networks, vol. 2, pp. 302-309, no. 2, March 1991.

[24] Ravi K. V. R. and R. Palaniappan Neural Network Classification of Late gamma band electroencephalogram features Soft Comput, vol. 10, no. 2, pp.163–169, 2006.

[25] Haykin S. Neural Network- A Comprehensive foundation, 2nd Edition, Prentice Hall, 2000.

[26] Hagen M., H. Demuth, and M. beale, Neural Network design Boston MA., PWS Publishing, 1996.

[27] Zhou Shang– Ming, John Q.Gan, Francisco Sepulved Classifying mental tasks based on features of higher-order statistics from EEG signals in brain-computer interface Information Sciences: an International Journal, Volume 178 ,Issue 6, Pages 1629-1640,2008.

[28] Cheng M., Xiarong Gao, S.Gao, D.Xu Design and implementation of a brain computer interface with high transfer rates IEEE, Transactions on Biomedical Engg., vol-49, no. 10 October 2002.

[29] Ting Wu , Yan Guo-zhenga, Yang Bang-huaa, Sun Hong EEG feature extraction based on wavelet packet decomposition for brain computer interface Measurement, Elsevier journal 41 2008, 618–625, 2008.

[30] Nikolaev A.R. and A.P. Anokhin EEG frequency ranges during reception and mental rotation of two and three dimensional objects. Neuroscience and Bheaviour physiology , v.-28, no-6, 1998.

[31] Osaka M. Peak alpha frequency of EEG during a mental task: task difficulty and hemisphere difference Psychophysiology,pp 101-105,vol.21,1984.

V. BIOGRAPHIES

Vijay Khare is currently pursuing his PhD in Bio Signal Processing at the Indian Institute of Technology, Delhi. He did his M.Tech in Instrumentation & Control, from NSIT Delhi. He is currently, with the Dept. Electronics and Communications Engineering at the Jaypee Institute of Information Technology. His research interests are Neural Networks, Brain Computer Interfacing, and Control Systems.

Dr.Jayashree Santhosh completed her B.Tech in Electrical Engineering from University of Kerala, M Tech in Computer & Information Sciences from Cochin University of Science and Technology, Kerala and Ph.D from IIT Delhi. She is a Fellow member of IETE, Life member of Indian Association of Medical Informatics (IAMI) and Indian Society of Biomechanics (ISB). Her research interests include IT in Healthcare Systems and was associated with a project on IT in Health Care at City University of Hong Kong. She is also associated with various projects with Centre for Bio-Medical Engineering at IIT Delhi in the area of Technology in Healthcare. Her research interests focus on Brain Computer Interface Systems for the Handicapped and in Neuroscience.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Prof. Sneh Anand is a professor and head, Center for Biomedical Engineering, Indian Institute of Technology, Delhi. She did B.Tech in Electrical Engg, from Punjab University, Patiala, and M.Tech in Instrumentation & Control from IIT Delhi and Ph.D. in Biomedical Engg. from IIT Delhi. Her research interests include biomedical instrumentation, rehabilitation engineering, biomedical transducers and Sensors.

Dr. Manvir Bhatia is the Chairperson of Dept. of SleepMedicine at Sir Ganga Ram Hospital, New Delhi and is also a Senior Consultant Neurologist.Dr. Manvir Bhatia completed her MBBS in 1981, and Doctor of Medicine in 1986 from Christian Medical College and Hospital, Ludhiana. DM in Neurology 1993, from All India Institute of Medical Sciences.She is a member of Indian Academy of Neurology, Indian Epilepsy Society, Indian Sleep Disorders Association, World Association of Sleep Medicine, International Restless Legs Society Study Group and American Academy of Electrodiagnostic Medicine.Dr. Manvir Bhatia has been invited to deliver lectures in National & International workshops, conferences on topics related to Neurology, Epilepsy, Sleep Medicine and has sleep published papers in leading journals.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Sixth order Butterworth Characteristics using LV MOCCII and Grounded Components.

T. Parveen
Electronics Engineering Department, Z. H. College of Engineering & Technology
AMU, Aligarh, INDIA
tahiraparveen2@gmail.com

***Abstract*— This paper introduces an active realization of the sixth order current mode Butterworth filter function using low voltage(LV) Multioutput current conveyors (MOCCII) and grounded passive components. The proposed realization is based on cascading an insensitive single input multi output (SIMO) current mode universal biquadratic filter (UBF). The UBF is constructed employing only two MOCCIIs, four grounded components, that lead to simple structure, easy to design and suitable for IC fabrication. The proposed UBF can realize all standard biquadratic responses without any matching conditions and has current outputs at a high impedance terminal, which enable easy cascadablity. An example of eighth order current mode Butterworth filter has been considered. The filter has the advantages of minimum requirement of active and passive component count, low sensitivity, and high performance. The performance of the filter is verified through PSPICE simulation using low supply voltage.**

Keywords-component; Current mode circuits, High order Butterworth filter, Multioutput current conveyors (MOCCIIs), Universal biquadratic filter.

I. INTRODUCTION

The current mode signal processing techniques have been received a wide attention due to its wide band width, improved linearity, wide dynamic range, low voltage operation as compared to voltage mode signal processing. By introducing Multi output current conveyors (MOCCIIs) in active filter designs, new more advantageous topologies [5] have been obtained, by which current outputs and current feedback can be developed when multiple current outputs are used [7].

Recently design of current mode universal biquadratic filters (UBF) with single input multi output (SIMO) have received considerable attention due to their convenience, high performance and greater functional versatility in terms of signal processing for practical applications [1-7]. Interconnections of relevant output currents provides the low pass, band pass, high pass, band elimination and all pass filter responses from the same circuit.

This paper presents a minimum active-RC circuit realization of eighth order current mode Butterworth filter function using low voltage Multioutput current conveyors **(**MOCCII) and grounded Passive components. The proposed realization is based on cascading a novel single input multi output (SIMO) current mode universal biquadratic filter (UBF), which uses two Multi output current conveyors (MOCCII) along with only four grounded passive components. The presented UBF can realize the low pass, high pass, band pass, band elimination and all pass filter responses through appropriate selection of input and output terminals without any matching conditions. The filter circuit is simple in structure and has high impedance outputs enables easy cascading in current mode operations, and can be used for the realization of any type of high order filter function. The proposed UBF circuit also has the additional advantage of low component count, and high performance at low supply voltage, over previously reported literature [1-7, 10]. The eighth order current mode realization with Butterworth coefficients has not yet been reported. The design may be extended for maximally flat Butterworth characteristics by cascading another MOCCII UBF stages.

II. REALIZATION OF UNIVERSAL BIQUADRATIC FILTER:

The presented universal biquadratic filter uses only two MOCCIIs along with two grounded capacitors and two grounded resistors as shown in Figure 2. The Multi Output Current Conveyors (MOCCIIs) is shown in Figure 1. It is characterized by

$$i_y = 0 \quad v_x = v_y \quad i_{zi}{}^{+} = +i_x \quad i_{zi}{}^{-} = -i_x \tag{1}$$

where i = 1, 2, 3…

Analysis of the circuit yields the following current transfer functions.

$$T_{LP}(s) = \frac{I_{LP}}{I_{IN}} = -\frac{\frac{1}{R_1 R_2 C_1 C_2}}{D(s)} \tag{2}$$

$$T_{BP}(s) = \frac{I_{BP}}{I_{IN}} = \frac{-\frac{s}{R_1 C_1}}{D(s)} \tag{3}$$

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

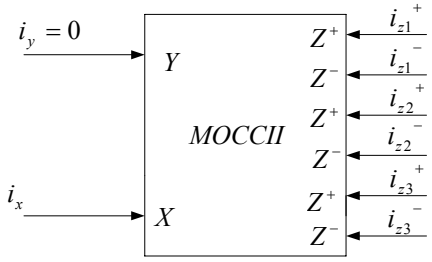


Figure 1. Symbol of MOCCII

$$T_{BE}(s) = \frac{I_{BE}}{I_{IIN}} = \frac{s^2 + \dfrac{1}{R_1 R_2 C_1 C_2}}{D(s)} \tag{4}$$

where, the denominator is given by:

$$D(s) = s^2 + \frac{s}{R_1 C_1} + \frac{1}{R_1 R_2 C_1 C_2} \tag{5}$$

From eqn. (2), (3) and (4) it is seen that inverting low pass, inverting band pass and non- inverting band elimination filters are realized at three outputs. Non- inverting high pass filter is realized just by connecting high impedance outputs I_{LP} and $I_{BE.}$ And all pass filter is realized by connecting the high impedance outputs I_{BP} and $I_{BE.}$ The realized high pass and all pass filter responses respectively are given by the following equations.

$$T_{HP}(s) = \frac{I_{HP}}{I_{IIN}} = \frac{s^2}{D(s)} \tag{6}$$

$$T_{AP}(s) = \frac{I_{AP}}{I_{IIN}} = \frac{s^2 - \dfrac{s}{R_1 C_1} + \dfrac{1}{R_1 R_2 C_1 C_2}}{D(s)} \tag{7}$$

The pole frequency ω_o and the quality factor Q of the filters are given by

$$\omega_o = \sqrt{\frac{1}{R_1 R_2 C_1 C_2}} \qquad Q = \sqrt{\frac{R_1 C_1}{R_2 C_2}} \tag{8}$$

III. Realization of Butterworth higher order filters:

Here we consider the realization of 6th order Butterworth low pass filter by cascading three sections of current mode universal filter biquadratic filter (UBF) employing MOCCIIs shown in Figure 2.

The normalized Butterworth transfer function for the resulting 6th order CM low pass filter is given by [9]

$$T(s) = \frac{1}{(s^2 + 1.932s + 1)(s^2 + 1.414s + 1)(s^2 + 0.518s + 1)} \tag{9}$$

The normalized pole frequency is at $\omega_{o=1}$.

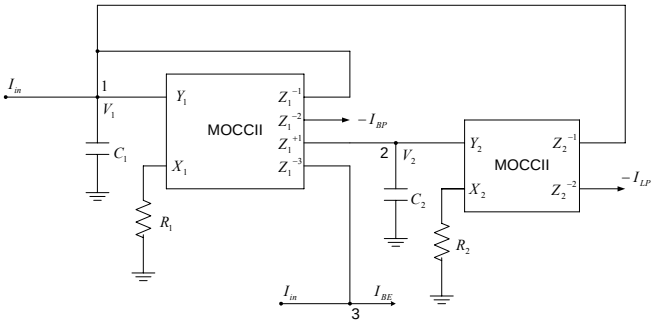


Figure 2. Cascading stage (CM MOCCII UBF) for sixth order low pass Butterworth filter

The transfer function can be de-normalized by replacing S → s/ω_o to give the required sixth order filter function at given pole- ω_o and pole-Q. The pole-Q of an individual biquadratic filter section is simply the reciprocal of the coefficients of s in eqn. (9) [8]. The values of pole-Q are 0.518, 0.707, and 1.932, respectively, for the three UBF to be cascaded. The UBF can be designed using these values of Q and for a given pole frequency. It is seen that no additional buffers are employed in the realization. The filter's pole frequency ω_o and the quality factor Q of the current mode MO-CCII-based UBF are given by:

$$\omega_o = \sqrt{\frac{1}{R_1 R_2 C_1 C_2}} \qquad Q = \sqrt{\frac{R_1 C_1}{R_2 C_2}} \tag{10}$$

IV. Non Ideal Effects

Taking the non-idealities α_i and β_i, i = 1 and 2, into consideration, the current transfer functions of the UBF are given then by:

$$T_{LP}(s) = \frac{I_{LP}}{I_{IN}} = -\frac{\dfrac{\alpha_1 \alpha_2 \beta_1 \beta_2}{R_1 R_2 C_1 C_2}}{D'(s)} \tag{11}$$

$$T_{BP}(s) = \frac{I_{BP}}{I_{IN}} = \frac{-\dfrac{s\alpha_1 \beta_1}{R_1 C_1}}{D'(s)} \tag{12}$$

$$T_{BE}(s) = \frac{I_{BE}}{I_{IIN}} = \frac{s^2 + \dfrac{\alpha_1 \alpha_2 \beta_1 \beta_2}{R_1 R_2 C_1 C_2}}{D'(s)} \tag{13}$$

$$T_{HP}(s) = \frac{I_{HP}}{I_{IIN}} = \frac{s^2}{D'(s)} \tag{14}$$

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

$$T_{AP}(s) = \frac{I_{AP}}{I_{IIN}} = \frac{s^2 - \frac{s\alpha_1\beta_1}{R_1C_1} + \frac{\alpha_1\alpha_2\beta_1\beta_2}{R_1R_2C_1C_2}}{D'(s)} \quad (15)$$

and

$$D'(s) = s^2 + \frac{s\alpha_1\beta_1}{R_1C_1} + \frac{\alpha_1\alpha_2\beta_1\beta_2}{R_1R_2C_1C_2} \quad (16)$$

where the pole frequency ($\omega_{o'}$) and the quality factor (Q′) of the filters obtained from D′(s) are given by:

$$\omega_o' = \sqrt{\frac{\alpha_1\alpha_2\beta_1\beta_2}{R_1R_2C_1C_2}}, \quad Q' = \sqrt{\frac{R_1C_1\alpha_2\beta_2}{R_2C_2\alpha_1\beta_1}} \quad (17)$$

At low to medium frequencies ($f \leq 10$ MHz), the circuit continues to provide standard second order responses. The pole- ω_o is slightly lowered, but the pole-Q remains unaffected by the non-idealities.

V. Sensitivity Study

The sensitivity of filter parameters are evaluated with respect to active and passive elements and are given below.

$$S^{\omega_o}_{R_1,R_2,C_1,C_2} = -\frac{1}{2} \qquad S^{Q}_{R_1,C_1,\alpha_2,\beta_2} = \frac{1}{2}$$

$$S^{\omega_o}_{\alpha_1,\alpha_2,\beta_1,\beta_2} = \frac{1}{2} \qquad S^{Q}_{R_2,C_2,\alpha_1,\beta_1} = -\frac{1}{2} \quad (18)$$

From the above calculation it is evident that the sensitivities of ω_o and Q with respect to passive components are not more than half in magnitude. This shows the attractive sensitivity feature of the UBF.

VI. Design and Simulation

To demonstrate the performance of CM universal biquadratic filter , the circuit is simulated using PSPICE level 3 parameters in 0.5µm CMOS process with supply voltages V_{DD} = -V_{SS} = 0.75V, using MOCCII+ model derived from CCII [9]is shown in Figure 3. The dimensions of the MOS transistors of CCII are listed in Table 1.

TABLE I. MOSFET Dimensions of the CCII.

MOSFET	W(μ m)	L(μ m)
M_1, M_2	25	0.5
M_3, M_9	66	1
M_4, M_6, M_7	4	0.5
M_5	12	0.5
M_8, M_{10}	45	1

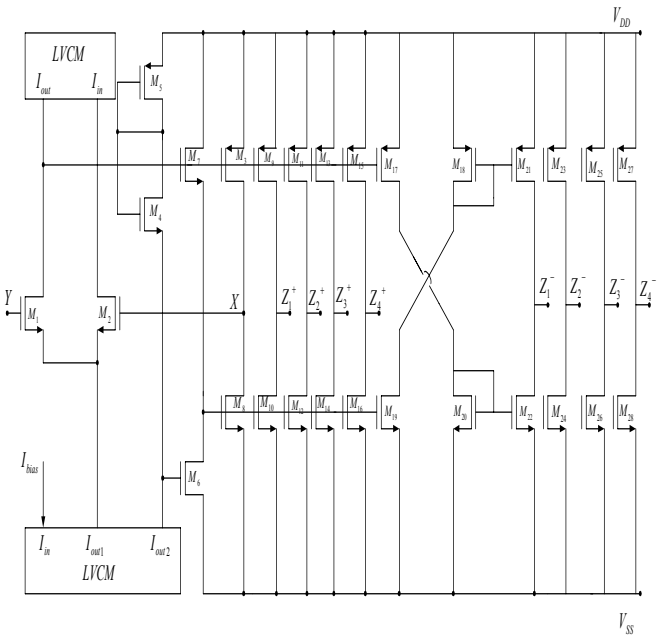


Figure 3. CMOS Circuit for MOCCII

Initially the UBF was designed for f_o = 1 MHz and Q = 0.707. For $C_1 = C_2 = 11$ pF, equation (8) yields $R_1 = 10$ KΩ, and $R_2 = 20$ KΩ. The simulated LP, BP, BE response of current mode UBF response is shown in Figure 4 shows good agreement with the theory.

The frequency tuning of the BPF response at Q = 5 is next investigated by changing the center frequency f_o of the band pass filter through resistor R_2. The BP response curves corresponding to f_o = 300 KHz, f_o = 500 KHz, and f_o = 1MHz are given in Figure 5, which exhibit good agreement with the theory.

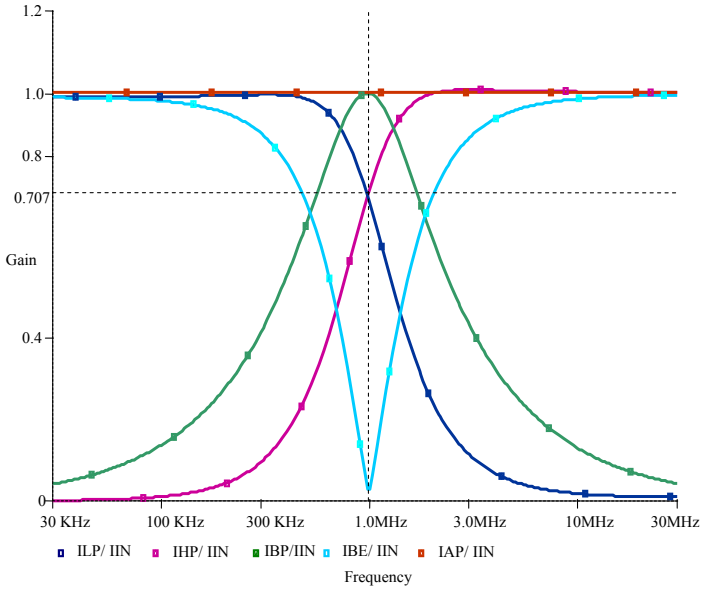


Figure 4. The simulated UBF response at f_o = 1 MHz

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

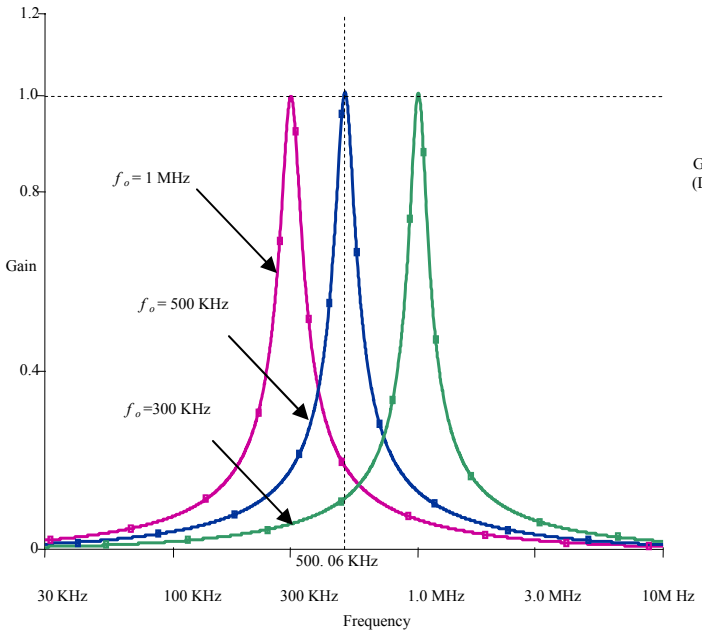


Figure 5. Frequency tuning of BPF at Q = 5

Next we presents an example for the realization of 6th order Butterworth low pass filter using low pass filter of UBF of Figure 2. To evaluate the performance of the sixth order Butterworth low pass filter, the circuit is designed for a pole frequency f_o =1MHz. The values of capacitors are selected equal for convenient in IC implementation and are assumed to be equal to 11pF. The resistors for each section are designed to satisfy the equation (8). The designed values for each section are given below.

Section-I: R_1 = 5.18 KΩ, R_2 = 19.32 KΩ, for pole Q = 0.518

Section-II: R_1 = 7.076 KΩ, R_2 = 14.15 KΩ, for pole Q = 0.707

Section-III: R_1 = 19.31 K KΩ, R_2 = 5.18 K KΩ, for pole Q = 1.932

It is seen that the simulated pole frequency of 1.04 MHz, is obtained from the simulation, which verifies the design. Figure 6 gives the stop band attenuation of 120 DB/decade, verifying the 6th order low pass response. Through the entire range, the simulated and theoretical responses overlap, showing close agreement with theory.

The proposed circuit can also be used to realize other higher order responses, such as, band pass, high pass and band elimination filters, through a simple electronic switching arrangement, for selecting the desired response of the UBF. The result of the 6th order band pass filter is shown in Figure 7, with a simulated pole frequency f_o =1.02 MHz and a pole-Q of 1.84. The slopes below f_o and above f_o are each 60 DB/decade, thus verifying the 6th order band pass response. At the pole frequency f_o =1 MHz, the gain is equal to unity.

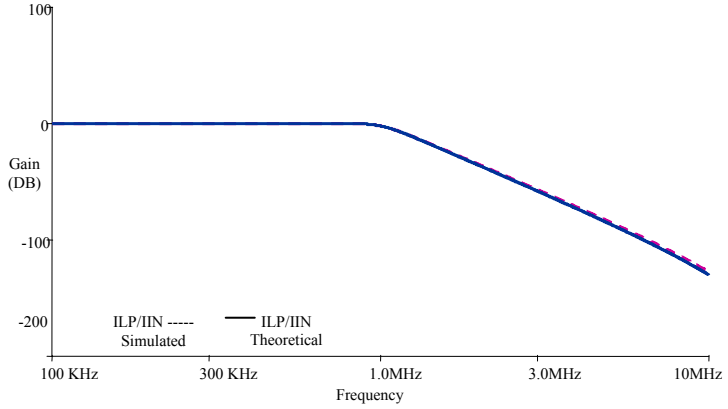


Figure 6. Frequency response of sixth order CM LPF in DB

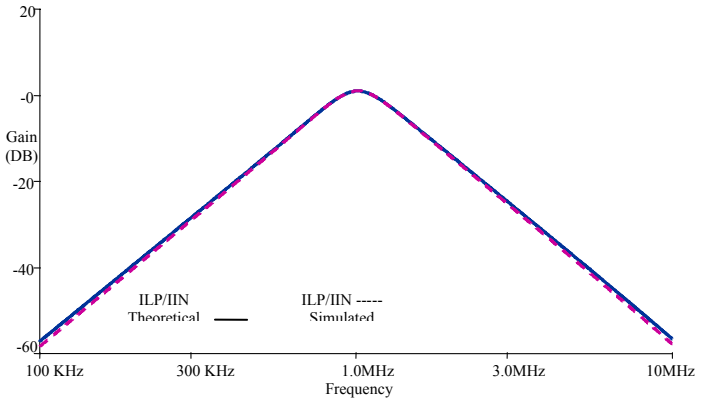


Figure7. Frequency response of sixth order CM BPF in DB

VII. CONCLUSION:

The MOCCII based current mode universal biquadratic filter is used to realize sixth order Butterworth low pass filter by cascading biquadratic filter sections of UBF of Fig. 2, without using any additional current followers. With this cascade approach the realization of higher order filter is reduced to a much simpler realization of only second order filters. The proposed circuit uses grounded capacitors hence is suitable for IC implementation. The proposed circuit also has low sensitivity, low component count, at low supply voltage of ± 0.75V.

REFERENCES

[1] S. Ozoguz and C. Acar, "Universal current mode filter with reduced number of active and passive elements", Electron. Lett., Vol.33, pp.948-949, 1997.

[2] A.M. Soliman,"Current Conveyor filters: classification and review", Microelectronics. J., Vol.29, pp. 133-149, 1998.

[3] A. Toker, and S. Ozoguz, "Insensitive current mode universal filter using dual output current conveyors,"Int. J. Electronics, Vol 87, No.6 pp.667-674, 2000.

[4] H. Y. Wang and C. T. Lee, "Versatile insensitive current mode universal filter implementation using current conveyors", IEEE Trans. Circuits Sys.II:, Analog and Digital Signal Processing, Vol.48, No.4, pp. 411-413, 2001.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

[5] E. Yuce, S. Minaei and O. Cicekoglu," Universal current mode active-C filter employing minimum number of passive elements", Analog Integ. Circ. Sig. Process., Vol.46, pp. 169-171, 2006.

[6] J. W. Horng, C. L. Hou, C. M. Chang, J. Y. Shie and C. H. Chang," Universal current filter with single input and three outputs using MOCCIIs, Int. J. Electronics, Vol.94, No.4, pp.327-333, 2007.

[7] N. Herencsar and K. Vrba, "Tunable Current-Mode Multifunction Filter Using Universal Current Conveyors", Third International Conference on Systems, ICONS 08, pp. 1-6, 2008.

[8] R. Schaumann, M. E. Van Valkenburg, "Design of analog filters", Oxford University press, 2003

[9] Rajput, S. S., Jamuar, S. S., "Advanced Applications of Current Conveyors: A Tutorial**",** J. of Active and passive Electronic Devices, Vol.2, pp.143-164, 2007.

[10] M. Kumar, M.C. Srivastava, U. Kumar, "Current Conveyor Based Multifunction Filter", International Journal of Computer Science and Information Security, Vol. 7, No. 2, pp.104-107, 2010.

[11] S.S. Rajput, and S.S. Jamuar, "A current mirror for low voltage, high performance analog Circuits", Analog Integrated Circuits and Signal Processing, Vol.36, pp 221-233, 2003.

AUTHORS PROFILE

Tahira Parveen received the B.Sc. Engineering and M.Sc. Engineering degrees from Z.H.College of Engineering & Technology, A.M.U. Aligarh, India in 1984 and 1987 respectively. She has obtained her Ph.D degree in 2009 from Electronics Engineering Department, Z.H.College of Engineering & Technology, A.M.U., Aligarh, India. She is currently an Associate professor in the department of Electronics Engineering, Z.H.College of Engineering & Technology, A.M.U. Aligarh, India. She has over 24 years of teaching and research experience. Her research interests are Electronic circuits and system design, Analog filters and Analog signal processing. She has published over 24 research papers in National and International Journals.

She has reviewed a book on "Basic Electronics" of McGraw-Hill Publisher. She is a Fellow member of IETE (India). She received Vijay Rattan Award for outstanding services achievements & contributions. Her biography is published in Marquis Who's Who in the World (USA) 2007. She is Reviewer for the Journal: International Journal of Circuit Theory and Applications, National and International conferences. Her book entitled" A Textbook of Operational Transconductance Amplifier & Analog Integrated Circuits" is published by I.K. International publisher in 2009.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

A Lightweight Secure Trust-based Localization Scheme for Wireless Sensor Networks

P. Pandarinath
Associate Professor, CSE, Sir C R Reddy College of Engineering
Eluru-534001, Andhra Pradesh
pandarinathphd@gmail.com
sriram310@gmail.com

M. Shashi
Head of the Department, Department of CS&SE, Andhra University,
Visakhapatnam- 530 003,
Andhra Pradesh
smogalla2000@yahoo.com

Allam Appa Rao
Vice Chancellor,
JNTU Kakinada,
Kakinada,
Andhra Pradesh
apparaoallam@gmail.com

***Abstract*—Location based security plays an important role in the trustworthiness of wireless sensor networks and the results that are obtained from them. Enforcement of location-aware security policies requires trusted location information. As more of these location-dependent services get deployed, the mechanisms that provide location information will become the target of misuse and attacks. In this paper, we propose to design a protocol that validates the reliability of location information associated with event reports. The protocol depends on the collaborative interaction of the network nodes to find compromised nodes. Nodes in the network record information of routing paths taken by packets through the network. Upon receiving the route request packets from the nodes, the sink checks whether their route matches a historically expected behavior by packets from the same claimed location. A trust value for the location claim of this request is then created by the sink. The attached trust values will be used to certify the truthfulness of the packets location information. Since this scheme does not involve any complex cryptographic operations, it has less overhead and delay. By simulation results, we show that our proposed scheme attains good delivery ratio with reduced delay and overhead.**

Keywords- Localization; Wireless Sensor Networks; Trust; Security issues; Lightweight Secure Trust-based Localization (LSTL)

I. INTRODUCTION

A. Sensor networks and its applications

Wireless Sensor Networks (WSNs) are a specific kind of ad hoc networks, highly decentralized, and without infrastructure. They are building up by deploying multiple micro transceivers, also called sensor nodes that allow end users to gather and transmit environmental data from areas which might be inaccessible or hostile to human beings. The transmission of data is done independently by each node, using a wireless medium. The energy of each node is limited to the capacity of its battery. The consumption of energy for both communication and information processing must be minimized [1].

Wireless sensor networks are an area of great interest to both academia and industry. They open the door to a large number of military, industrial, scientific, civilian and commercial applications. They allow cost-effective sensing especially in applications where human observation or traditional sensors would be undesirable, inefficient, expensive, or dangerous [2].

B. Localization in Sensor Networks

In sensor networks, without earlier knowledge of their location, nodes are organized into an unintentional infrastructure dynamically. Localization or position estimation problem refers to the problem of finding the positions of all the nodes provided a few location aware nodes, relative distance and angle information between the nodes. The essential and vital problem in wireless sensor network operation is to determine the physical positions of sensors for following reasons.

- It is always necessary to have the position information of sensors attached, in order to use the data collected by the sensors. For instance, in sensor networks, the physical location of each sensor should be known in advance for identifying the position of the detected objects in order to detect and track objects.
- With the knowledge of the geographic positions of sensors, many communication protocols of sensor networks are built. In majority of cases, there is no supporting infrastructure available to locate the sensors as they are deployed devoid of their position information known in advance.

Hence it is necessary to find the position of each sensor in wireless sensor networks after deployment [3].

C. Phases of Localization

The localization phase is a very critical step that must be secured in order to ensure the integrity of the WSN and its associated services.

- Firstly, this process allows the sensor to set up the necessary parameters to establish the paths that will lead their data towards end users.
- The knowledge of their position is also an essential prerequisite for the final application that processes the data collected by sensors, i.e., the user needs to know the origin of collected data before using it.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

- Finally, the end users might want to query some nodes by sending the position where information needs to be collected [1].

Many localization schemes have been proposed for sensor networks in recent years without depending on expensive GPS devices. The majority of existing schemes assume some unique nodes, called beacon nodes, which have the potential to know their own locations through either GPS receivers or manual configuration. The remaining nodes or the non-beacon sensor nodes can be equipped with somewhat cheap measuring devices for directionality, signal strength, or time of arrival, etc. The non-beacon nodes find their own locations using these measurements and the locations of three or more beacon nodes. This method is known as the beacon-based technique, which involves two stages in location discovery.

i. A sensor node calculates its distances to each neighbor of the node using the received signal information.

ii. With all these distance calculations, sensor nodes calculate the actual location of the node. [4]

D. Security Issues on Localization

An important concern for various applications of WSNs is the ability to validate the integrity of the sensor network as well as the retrieved data. Various types of security attacks include

- The injection of false information into the regular data stream,
- The alteration of routing paths due to malicious nodes advertising false positions (sink holes and worm holes), and
- The forging of multiple identities by the same malicious node.

Thus, location based security plays an important role in the trustworthiness of WSNs and the results that are obtained from them [5].

Enforcement of location-aware security policies requires trusted location information. As more of these location-dependent services get deployed, the mechanisms that provide location information will become the target of misuse and attacks. In particular, the location infrastructure will be affected by many localization-specific threats that cannot be tackled through traditional security services. Therefore, as we move forward with deploying wireless systems that support location services, it is sensible to integrate appropriate mechanisms that protect localization techniques from these new forms of attack [6].

The wormhole attack is a typical kind of secure attacks in WSNs. It is launched by two colluding external attackers which do not authenticate themselves as legitimate network nodes to other network nodes. One of the wormhole attackers overhears packets at one point in the network, tunnels them through the wormhole link to another point in the network, and the other wormhole attacker broadcasts the packets among its neighborhood nodes. This may cause a severe impact on the routing and localization procedures in WSNs [7].

E. Problems Identified and Proposed Scheme

The proposed solutions to mitigate some of these localization attacks always involve traditional security techniques. But, it is unlikely that traditional security will be able to remove all threats to wireless localization. We therefore consider that instead of providing solutions for each attack, it is essential to achieve robustness to unpredicted and non-filterable attacks. Particularly, localization must function properly even in the presence of these attacks [6].

Digital signatures can prevent bogus seeds from injecting bogus location messages by authenticating seeds' transmissions to nodes. This could be done by distributing public keys corresponding to the seeds' private keys to each node before deployment. But public key encryption operations are often too computationally expensive for sensor nodes, however, and the long messages required drain power resources [8].

Another approach would be to use the mTesla protocol, by preloading each node with the initial hash chain value and each seed with the initial secret. This would save the expense of public key operations, but would delay localization until the next key in the hash chain is released. It would also require loose synchronization among the seeds [8].

Moreover, since distance measurements are susceptible to distance enlargement/reduction, such techniques may not be used to infer the sensor location [9].

In this paper, we propose to design a protocol that validates the reliability of location information associated with event reports. The protocol depends on the collaborative interaction of the network nodes to find compromised nodes. Each active node automatically has some knowledge of the activity within the network which can be used in determining the anomalous behavior. Nodes in the network record information of routing paths taken by packets through the network. Upon receiving a packet, nodes check whether their route matches a historically expected behavior by packets from the same claimed location. A trust value for the location claim of this packet is then created and propagated to the sink. The attached trust values will be used by the sink to certify the truthfulness of the packets location information.

II. Related Work

J. G. Alfaro, M. Barbeau and E. Kranakis [1] have provided three algorithms that enable the sensor nodes of a Wireless Sensor Network to determine their location in presence of neighbor sensors that may lie about their position. Their algorithms minimize the number of trusted nodes required by regular nodes to complete their process of localization. Also their algorithms always work for a given number of neighbors provided that the number of liars is below a certain threshold value, which is also determined. The three algorithms that they have presented guaranteed that regular nodes in the WSN always obtain their position provided that the number of liars in the neighborhood of each regular node is below a certain threshold value, which they determine for each algorithm. Their three algorithms allow the regular nodes to identify and isolate nodes that are providing false information about their position. Moreover, their

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

algorithms minimize the necessary number of trusted nodes required by regular sensors to complete their process of localization. They also guarantee a small exchange of data between nodes, minimizing in this manner the impact that the localization process has in terms of energy and battery life of sensors.

Kaiqi Xiong and David Thuente [4] have proposed novel schemes for secure dynamic localization in sensor networks. Their proposed schemes can tolerate up to 50% of beacon nodes being malicious, and they have linear computation time with respect to the number of reference nodes. They also showed that their schemes are applicable and resilient to attacks from adversaries. They proposed several methods for secure location discovery in sensor networks which are developed through the technique of beacon-based localization.

Honglong Chen et al [7] have investigated the impact of the wormhole attack on the localization and they proposed a novel consistency-based secure localization scheme against wormhole attacks, which includes wormhole attack detection, valid locator's identification and self-localization. They developed an enhanced identification approach which obtains better performance than the basic identification approach.

Yanchao Zhang et al [10] have first analyzed the security of existing localization techniques. They then developed a mobility-assisted secure localization scheme for UWB sensor networks. They didn't intend to provide brand-new localization techniques for UWB sensor networks. Instead, they focused on analyzing and enhancing the security of existing approaches when applied in adversarial settings. In addition, they have proposed a location-based scheme to enable secure authentication in UWB sensor networks.

Srdjan C apkun et al [11] have proposed a secure localization scheme for sensor networks based on the received signal strength (RSS) ranging techniques. Their scheme enables the network authority to obtain locations of sensor nodes in the presence of an attacker. Also their proposed scheme uses a small number of anchor nodes with known locations that provide points of reference from which the sensors locations are computed. Their scheme also makes use of robust localization and time synchronization primitives which, appropriately combined, enable the detection of attacks on localization, within a realistic attacker model.

III. Proposed Lightweight Secure Trust-based Localization (LSTL) Scheme

A. System Design and Overview

Sensors will start generating event reports corresponding to their respective location, once the network becomes activated. A malicious sensor might then try to generate illegitimate event reports for locations other than its own. This can be detected when the nodes along the path from the source to destination, attaches trust values to passing data packets, ensuring the correctness probability of the declared source locations. These trust values evaluated based on gathered past traffic patterns combined with the claimed source location. When receiving packets, if any of the routing information diverges from projected traffic patterns, then the nodes have the chance to decrease the trust values associated with these packets. These decreased trust values replicate the appearance of an attacker in the routing pattern.

We assume that the network at the start is free from attackers for a short period of time. The novelty in this scheme lies in the efficient way of shortening the history of traffic pattern and capability of using this history to authenticate the accuracy of future packets.

As a part of the routing protocol, each sensor node will maintain a history and normalized count of each previously seen source destination pair for routed packets. New incoming packets from rarely seen sources will then be considered more suspicious and associated with a low trust value.

B. Distance Estimation

We first describe the meaningful way of comparing packet routes efficiently. Based on the sequence of nodes a packet has visited, we define a distance metric to measure the distance between the two paths. The distance is designed such that fake claimed locations for packets will result in large distances between real and expected paths. There are many generic ways to measure the distance between two curves in space.

Given a path R, we take k samples $\{R_1, R_2, \cdots, R_k\}$ on R. we define the distance between two paths R, $\overline{R}$ as the sum of squared distance between corresponding sample points.

$$\prod(R, \overline{R}) = \sum_{i=1}^{k} \| R_i - \overline{R}_i \|^2 \quad (1)$$

C. Trust Based Incentive Scheme

An additional data structure called Neighbor's Trust Counter Table (NTT) is maintained by each network node. Let $\{TC1, TC2, \cdots\}$ be the initial trust counters of the nodes $\{N_1, N_2 \cdots\}$ along the route from a source S to the sink D.

Since the node does not have any information about the reliability of its neighbors in the beginning, nodes can neither be fully trusted nor be fully distrusted. When a source S wants to send data to the sink D, it sends route request (RREQ) packets. It contains the source and destination ids and location of the source and a MAC computed over the accumulated path with a key shared by the sender and the destination

Each time, the sink estimates the distance between the two paths R and $\overline{R}$ towards the sink, using (1). If it is more than a maximum threshold Th1, then the trust counter value is decreased by a penalty of δ.

$$\text{(ie)} \qquad TC_i = TC_i - \delta, \quad (2)$$

Then the NTT of node Ni is modified with the values of TC_i. When the subsequent RREQ message reaches the destination, it checks the trust values of the intermediate nodes. The nodes are considered as well behaving nodes if the trust values are equal or greater than a trust threshold TC_{th}. On the other hand, the nodes are considered as misbehaving

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

nodes if the trust values are less than TC_{th} . Also nodes with trust values less than TC_{th} are prohibited from further transmissions.

D. Route Discovery Process

In the proposed protocol, once a node S want to send a packet to a sink D, it initiates the route discovery process by constructing a route request RREQ packet. It contains the source and destination ids and location of source and a MAC computed over the accumulated path with a key shared by each node.

When an intermediate node receives two RREQ packets from the same claimed locations, it retrieves the corresponding path information from their respective MAC and appends its id to the path and recreates the MAC with a key which is shared with the destination. It then forwards the RREQ to its neighbors.

Then the route request process is illustrated as below:

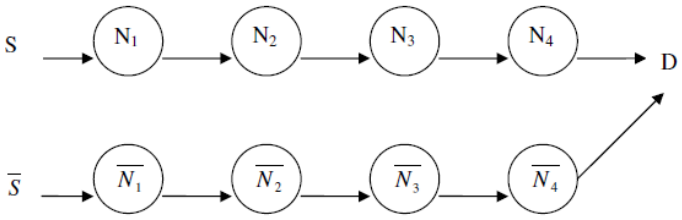


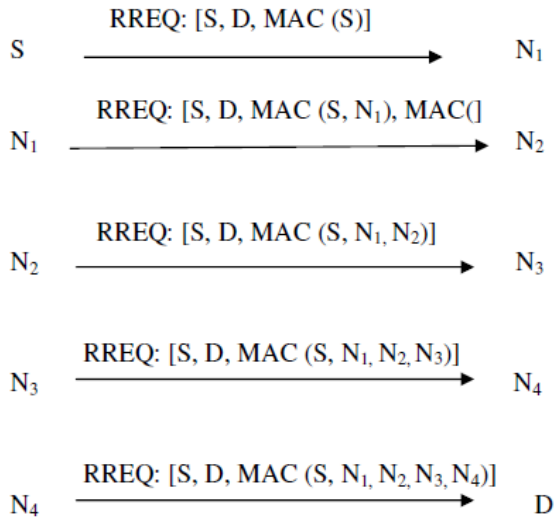
Figure1. Route Request Process

From the figure 1, we can see that there are two paths

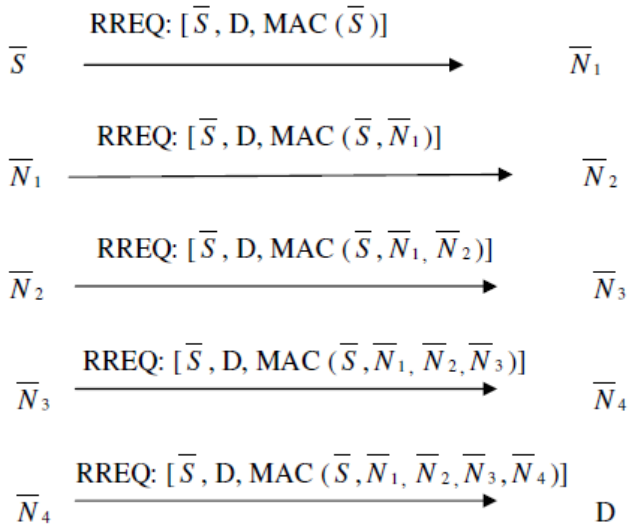
R: S ⟶ D and
$\overline{R}$: $\overline{S}$ ⟶ D

for the sink D.

For the path R , let $N_1, N_2, \cdots, N_4$ be the intermediate nodes between S and D. Then the route request process of path R is given by



For the path $\overline{R}$, let $\overline{N}_1$, $\overline{N}_2, \ldots \overline{N}_4$ be the intermediate nodes between $\overline{S}$ and D. Then the route request process of path $\overline{R}$ is given by



When RREQs of both R and $\overline{R}$ reaches the sink, from the received MAC values, it calculates

$V = D(N_i, \overline{Ni})$ where

$$D(N_i, \overline{Ni}) = \sum_{i=1}^{4} \| N_i - \overline{N}_i \|^2 \quad \text{by} \quad (1)$$

Then it checks the value of V , based on which the trust values are increemented or decreemented for the corresponding nodes.

If $V < th_1$ *then*

$$CCN_i = CCN_i + \delta,$$

Else

$$CCN_i = CCN_i - \delta,$$

End if

where th_1 is the minimum threshold value for V and δ is the scale factor for increement or decreement.

The process is repeated for various time intervals and finally the value of credit counter is checked,

If $CCN_i > th_2$ then
RREP is sent
Else
The source is considered malicious,
RREQ is discarded
End if.

Where th_2 is the minimum threshold value for CCN

In our scheme, only nodes which are stored in the current route need to perform these cryptographic computations. So the proposed protocol is efficient and more secure.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

IV. PERFORMANCE EVALUATION

A. Simulation Parameters

We evaluate our Light weight Secured Trust based Localization (LSTL) Algorithm through NS2 simulation. We use a bounded region of 1000 x 1000 sqm, in which we place nodes using a uniform distribution. We assign the power levels of the nodes such that the transmission range and the sensing range of the nodes are all 250 meters. In our simulation, the channel capacity of mobile hosts is set to the same value: 2 Mbps. We use the distributed coordination function (DCF) of IEEE 802.11 for wireless LANs as the MAC layer protocol. The simulated traffic is Constant Bit Rate (CBR).

The following table summarizes the simulation parameters used

TABLE I. SIMULATION PARAMETERS

No. of Nodes	25,50,75 and 100
Area Size	1000 X 1000
Mac	802.11
Simulation Time	50 sec
Traffic Source	CBR
Packet Size	512
Transmission Range	250
Routing Protocol	AODV
Speed	5
Mobility model	Random way point
Attackers	5, 10, 15, 20 and 25

B. Performance Metrics

We compare the performance of our proposed LSTL with the SeRLoc [11]. We evaluate mainly the performance according to the following metrics:

Average end-to-end delay: The end-to-end-delay is averaged over all surviving data packets from the sources to the destinations.

Control overhead: The control overhead is defined as the total number of control packets exchanged.

Estimation Error: It is the estimation error, which indicates how close the estimated location is to the actual location.

Average Packet Delivery Ratio: It is the ratio of the number .of packets received successfully and the total number of packets transmitted.

A. Based On Nodes

In order to test the scalability, the number of nodes is varied as 25, 50, 75 and 100.

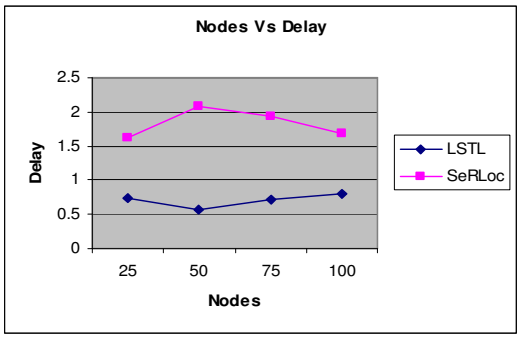


Figure2. Nodes Vs Delay

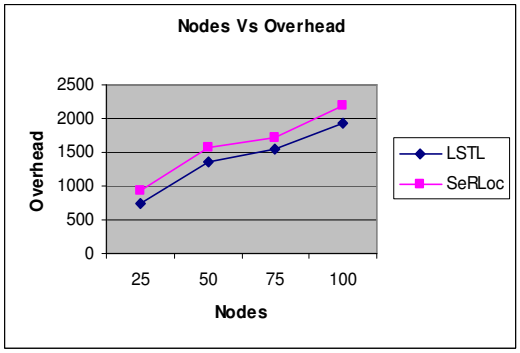


Figure3. Nodes Vs Overhead

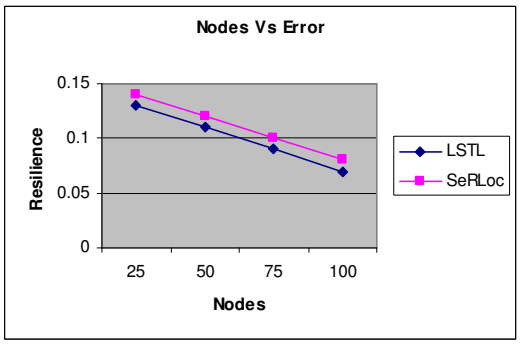


Figure4. Nodes Vs Error

Figure 2 show the end-to-end delay occurred for both LSTL and SeRLoc. As we can see from the figure, the delay is less for LSTL, when compared to SeRLoc.

Figure 3 shows the overhead for both LSTL and SeRLoc. As we can see from the figure, the overhead is less for LSTL, when compared to SeRLoc.

Figure 4 shows the error occurred for both LSTL and SeRLoc. As we can see from the figure, the error is less for LSTL, when compared to SeRLoc.

B. Based On Attackers

The number of attacker nodes is varied as 5, 10, 15, 20 and 25 in a 100 nodes scenario.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

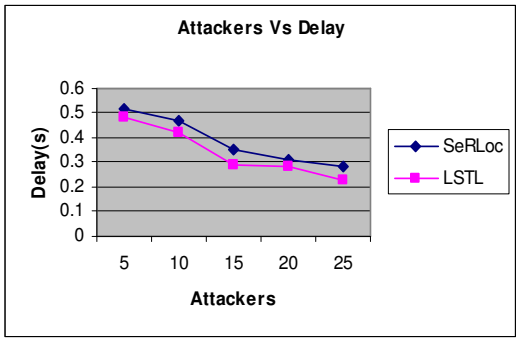


Figure5. Attackers Vs Delay

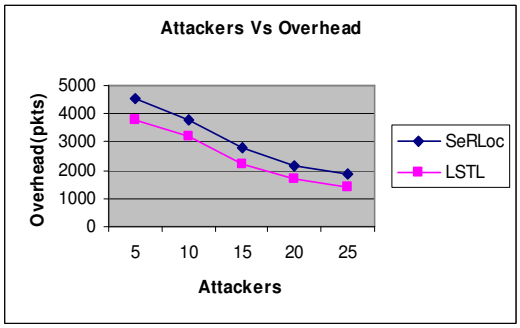


Figure6. Attackers Vs Overhead

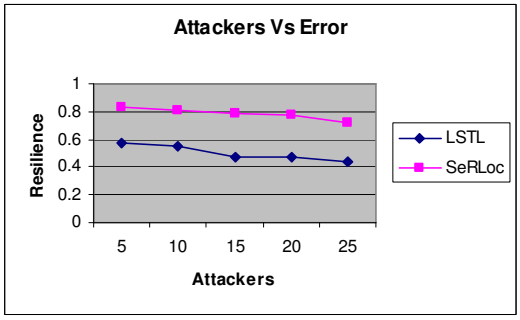


Figure7. Attackers Vs Error

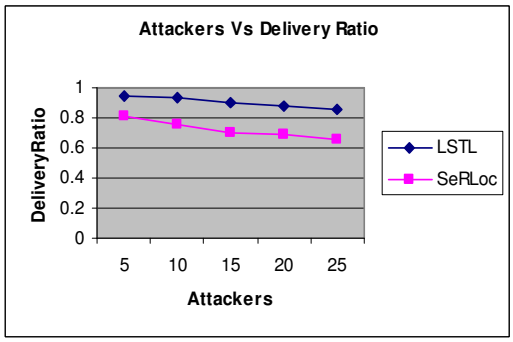


Figure 8. Attackers Vs DelRatio

Figure 5 show the end-to-end delay occurred for both LSTL and SeRLoc. As we can see from the figure, the delay is less for LSTL, when compared to SeRLoc.

Figure 6 shows the overhead for both LSTL and SeRLoc. As we can see from the figure, the overhead is less for LSTL, when compared to SeRLoc.

Figure 7 shows the error occurred for both LSTL and SeRLoc. As we can see from the figure, the error is less for LSTL, when compared to SeRLoc.

Figure 8 shows the delivery ratio for both LSTL and SeRLoc. As we can see from the figure, the delivery ratio is high for LSTL, when compared to SeRLoc.

V. CONCLUSION

In this paper, we have designed a protocol that validates the reliability of location information associated with event reports. The protocol depends on the collaborative interaction of the network nodes to find compromised nodes. When a source S wants to send data to the sink D, it sends route request (RREQ) packets. It contains the source and destination ids and location of the source and a MAC computed over the accumulated path with a key shared by the sender and the destination. When the RREQ packets reach the sink, it estimates the distance between the two paths R and $\overline{R}$ towards the sink. If it is more than a maximum threshold Th1, then the trust counter value is decreased by a penalty of δ. When the subsequent RREQ messages reach the destination, it checks the trust values of the intermediate nodes. The nodes are considered as well behaving nodes if the trust values are equal or greater than a trust threshold TC_{th}. On the other hand, the nodes are considered as misbehaving nodes if the trust values are less than TC_{th}. Also nodes with trust values less than TC_{th} are prohibited from further transmissions. Since this scheme does not involve any complex cryptographic operations, it has less overhead and delay. By simulation results, we have shown that our proposed scheme attains good delivery ratio with reduced delay and overhead.

REFERENCES

[1] J. G. Alfaro, M. Barbeau and E. Kranakis, "Secure Localization of Nodes in Wireless Sensor Networks with Limited Number of Truth Tellers", Proceedings of the Seventh Annual Communication Networks and Services Research Conference, 2009.

[2] Eric Sabbah et al, "An Application Driven Perspective on Wireless Sensor Network Security", Proceedings of the 2nd ACM international workshop on Quality of service & security for wireless and mobile networks, 2006.

[3] V. Vijayalakshmi et al, "Secure Localization Using Elliptic Curve Cryptography in Wireless Sensor Networks", IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.6, June 2008.

[4] Kaiqi Xiong and David Thuente, "Dynamic Localization Schemes in Malicious Sensor Networks", Journal of Networks, Vol. 4, no. 8, October 2009.

[5] E. Ekici et al, "Secure probabilistic location verification in randomly deployed wireless sensor networks", Ad Hoc Networks, Volume 6, Issue 2, April 2008.

[6] Zang Li et al, "Robust Statistical Methods for Securing Wireless Localization in Sensor Networks", Proceedings of the 4th international symposium on Information processing in sensor networks, 2005.

[7] Honglong Chen et al, "A Secure Localization Approach against Wormhole Attacks Using Distance Consistency", EURASIP Journal on Wireless Communications and Networking, Volume 2010 (2010).

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

[8] Lingxuan Hu and David Evans, “Localization for Mobile Sensor Networks”, Proceedings of the 10th annual international conference on Mobile computing and networking, 2004.

[9] Loukas Lazos and Radha Poovendran, ”SeRLoc: Robust Localization for Wireless Sensor Networks”, ACM Transactions on Sensor Networks (TOSN), 2005.

[10] Yanchao Zhang et al, “Secure Localization and Authentication in Ultra-Wideband Sensor Networks”, IEEE Journal on Selected areas in communications, 24(4), 829-835, 2006.

[11] Srdjan C apkun et al, “Secure RSS-based Localization in Sensor Networks”, Technical Reports 529, ETH Zürich, 09, 2006.

P. Pandarinath his B.Tech in CSE in 1994 and M.Tech in CSE in 2001. He is presently pursuing his doctorate. He has a total experience of 15 years. He is working as a Associate Professor in Sir C R Reddy college of Engineering. He has completed 80 projects and guided 100 projects. His areas of interests include Network Security and Cryptography, Advanced Computer Architecture, DataBase Management System, Computer Organization, Computer Networks and Bio-informatics.

Dr. M. Shashi completed her B.E. in EEE, and M. E. in Computers Engineering. She received Doctorate in Engineering. She is working as Head of the Department in Computer Science and Systems Engineering, Andhra University, Visakhapatnam. She has a total experience of 24 years. She presented 17 international journals and 3 national journals. She presented papers in 10 international conference and 8 national conferences. She attended 7 international and 25 national conferences. She guided 85 M.Tech projects. She received AICTE Career Award for Young Teachers in 1996-1998 at National Level and Best Ph.D. Thesis Prize for 1994 & 1995 at University Level. Her areas of specializations include Data Warehousing & Mining, AI, Data Structures, Soft computing and Machine Learning.

Dr. Allam Appa Rao received his B.Sc. (M.P.C.) in 1967 from Andhra University. He completed M.A. in Economics: Mathematical Economics and Econometrics, from Andhra University. He received the Ph.D. in Computer Engineering from Andhra University. in 1984. He is working as Vice Chancellor for JNTU Kakinada. He has presented 150 research articles. Dr Allam holds two patents as a co inventor for Method(s) of stabilizing and potentiating the actions and administration of brain-derived neurotrophic factor (BDNF) patent Number 20080234197 dated 25th September 2008 (Refer http://www.faqs.org/patents/inv/83318) and Method(s) of preventing, arresting, reversing and treatment of atherosclerosis Patent Number 20080279925 dated 13th November2008(Refer http://www.faqs.org/patents/inv/171637). He achieved best Researcher in Engineering in recognition of commendable record of research in Engineering, Andhra University, Visakhapatnam, India, 2003.

Mechanism to Prevent Disadvantageous Child Node Attachment in HiLOW

Lingeswari V Chandra, Kok-Soon Chai and Sureswaran Ramadass
National Advanced IPv6 Centre, Universiti Sains Malaysia
{ lingeswari , kschai, sures }@nav6.org

Gopinath Rao Sinniah
MIMOS Berhad, 57000 Kuala Lumpur
gopinath.rao@mimos.my

Abstract— **Vast research is being conducted in the area of Wireless Sensor Network in recent years due to it foreseen potential in solving problems covering many aspects of daily, industrial and ecological areas. One of the biggest challenges in the power and memory constrained sensor network is in establishing reliable network and communication among the nodes. . IP-based 6LoWPAN was introduced to give a new dimension to sensor network by enabling IPv6 to be applied to the wired as well as wireless sensors. An extendable and scalable Hierarchical Routing Protocol for 6LoWPAN (HiLOW) is one of three routing protocols which has been introduced specially for 6LoWPAN. HiLOW was designed by exploiting the dynamic 16 bit short addresses assignment capabilities featured by 6LoWPAN. HiLOW clearly defines the network setup process, address allocation method and routing mechanism. However there are shortcomings or issues pertaining HiLOW that make it less efficient. One of the major issues identified in HiLOW is in the process of selecting the parent node to attach with during the network tree setup. Disadvantageous parent selection could lead to significant shorter life span of the network which affects the reliability and stability of the network. In this paper we review the HiLOW routing protocol, highlight the issues revolving HiLOW and suggest a mechanism to prevent disadvantageous child node attachment in HiLOW. The proposed mechanism takes into consideration the LQI value, the potential parents' depth in the network and the average energy level of the parent in selecting the suitable parent node in order to provide a more reliable wireless sensor network.**

Keywords- 6LoWPAN, routing protocol, HiLOW, WSN, Hierarchichal routing protocol

I. INTRODUCTION

Wireless Sensor Network (WSN) is an area which is being vastly researched on in recent years due to its foreseen capability in solving many existing problems, ranging from day to day problems, industrial problems and up to ecological problems. WSN initially started as a military network where it was used to detect enemies, land mines and identifying own man. Now WSN usage has been extended to general engineering, agriculture monitoring, environmental monitoring, health monitoring and also home and office monitoring and automation.

The rapid growth and penetration of sensors to other areas are due to engineering contribution where more types of sensors has being developed and introduced while decreasing the size of sensor nodes, and power consumption and price of microprocessors while increasing the memory size of the nodes. Even though vast improvement has been witnessed from engineering perspective the nodes until today still faces the limitation in power and computational capacities and memory [1].

One of the most important elements after sensing activity but most energy costly element of the WSN is the communication part. Radio communication is typically the most energy consuming activities [2] and the reception energy is often as high as the transmission energy thus the network protocol introduced as well as the routing protocol needs to take into consideration the energy usage in setting up the network as well as complexity of computation during routing. A disadvantageous setup of network could lead to request of retransmission to occur and this would lead towards wastage of precious energy.

Many communication protocols have been introduced to WSN prior to the introduction of 6LoWPAN namely 802.15.1 Bluetooth [3], WirelessHart [4], ZWave [5], ZigBee [5] and others. Compared to the named communication network 6LoWPAN[6] was the first to introduce IPv6 to be applied to not only wireless but also wired sensor network. 6LoWPAN defines the network layer and also the transport layer and is able to be deployed on any sensors which are IEEE 802.15.4[7, 8] compliant. The 6LoWPAN stack is 30KB minimum in size which is smaller compared to the named protocols. The routing protocol for 6LoWPAN is an open area where it is open to introduction of new protocols. Till today there are three prominent routing protocols which have been introduced specifically for 6LoWPAN namely Hierarchical Routing Protocol (HiLOW) [10, 11], Dynamic MANET On-demand for 6LoWPAN (DYMO Low) [12] and 6LoWPAN Ad Hoc On-Demand Distance Vector Routing (LOAD) [13].

The remainder of this paper is organized as follows. Section 2 reviews the HiLOW protocol in detail and the issues in it and works done to improve HiLOW. Section 3 suggests a mechanism to avoid a disadvantageous parent child attachment during the routing tree set up. Section 4 presents the conclusion.

II. RELATED WORKS

A hierarchical routing protocol (HiLow) for 6LoWPAN was introduced by K. Kim in 2007 [10]. HiLOW is a routing protocol which exploits the dynamic 16-bits short address

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

assignment capabilities of 6LoWPAN. An assumption that the multi-hop routing occurs in the adaptation layer by using the 6LoWPAN Message Format is made in HiLOW. In the rest of this section we will discuss the operations in HiLOW ranging from the routing tree setup operation up to the route maintenance operation while highlighting issues revolving HiLOW. Other works done to solve some of the issues in HiLOW is also reviewed here.

A. HiLOW Routing Tree Setup, Issues and Works Done

The process of setting up the routing tree in HiLOW consists of a sequence of activities. The process is initiated by a node which tries to locate an existing 6LoWPAN network to join into. The new node will either use active or passive scanning technique to identify the existing 6LoWPAN network in its Personal Operation Space (POS).

If the new node identifies an existing 6LoWPAN it will then find a parent which takes it in as a child node and obtain a 16 bit short address from the parent. The parent will assign a 16 bit short address to a child by following the formula as in (1). An important element of HiLOW is that the Maximum Allowed Child (MC) need to be fixed for every network and all the nodes in the network is only able to accept child limited to the set MC. In the case where no 6LoWPAN network is discovered by the node then the node will initiate a new 6LoWPAN by becoming the coordinator and assign the short address by 0.

FC : Future Child Node's Address

MC : Maximum Allowed Child Node

N : Number of child node inclusive of the new node.

AP : Address of the Parent Node

$$FC = MC * AP + N \ (0 < N <= MC) \qquad (1)$$

Two potential issues have been identified in this process. First issue would be when the child node gets respond from more than one potential parent. There is no clear mechanism rolled out in selecting the suitable parent to attach with. If the new node chooses to join the first responding parent node, it could be bias to the parent as some parent might be burdened with more parents meanwhile other parents which is in the same level has less child or none at all. Selecting the parent based on first responded potential parent could also lead to fast depletion of energy to certain parent causing the life span of the network to be shorter and the stability to be jeopardized. Selection of parent without considering the link quality could cause towards high retransmission rate which will consume energy from the child node as well as parent node.

In [15] a mechanism to overcome the issue was suggested. Their mechanism suggests the potential parent node to provide the new child with its existing child node count (child_number). By issuing the child_number the node could select suitable parent which has less child nodes. The suggested mechanism performs well only when the potential parent node has same depth, same energy level and has different number of existing child. Their mechanism also does not take into consideration the quality of the link established between the parent node and child node. Therefore the suggested mechanism does not solve the arising issue completely. So in this paper we will be suggesting a mechanism to address this first issue by taking into consideration the link quality, the existing energy of the potential parent as well as the depth of the parent.

Second issue revolves around the MC value which is being fixed for all nodes. The current scenario works well if all the nodes have the same power conservation method; meaning if all the nodes in the network are either battery powered or mains-powered. In the case where some nodes are battery powered and the others are mains-powered, then this method is not advantageous. The main-powered nodes could support more nodes as their child node as they are affluent in energy. This is an open issue to be addressed in HiLOW, for present time the assumption that all nodes having same energy conservation have to be made. The activity of disseminating the MC value to joining nodes is also left in gray. This issue is not addressed in this paper.

B. Routing Operation in HiLOW

Sensor nodes in 6LoWPAN can distinguish each other and exchange packet after being assigned the 16 bits short address. HiLOW assumes that all the nodes know its own depth of the routing tree. The receiving intermediate nodes can identify the parent's node address through the defined formula (2). The '[]' symbol represents floor operation

AC : Address of Current Node

MC : Maximum Allowed Child

$$AP = [(AC-1) / MC] \qquad (2)$$

The receiving intermediate nodes can also identify whether it is either an ascendant node or a descendant node of the destination by using the above formula. When the node receives a packet, the next hop node to forward the packet will be calculated by the following three cases (3) which is defined in [10].No issues have been identified so far in this process.

SA : Set of Ascendant nodes of the destination node

SD : Set of Descendant nodes of the destination node

AA(D,k): The address of the ascendant node of depth D of the node k

DC : The depth of current node

C : The current node

Case 1: C is the member of SA (3)

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

The next hop node is AA (DC+1, D)

Case 2: C is the member of SD

The next hop node is AA (DC-1, C)

Case 3: Otherwise

The next hop node is AA (DC-1, C)

C. Route Maintenance in HiLOW

Each node in HiLOW maintains a neighbor table which contains the information of the parent and children node. When a node loses an association with its parent, it should to re-associate with its previous parent by utilizing the information in its neighbor table. In the case of the association with the parent node cannot be recovered due to situation such as parent nodes battery drained, nodes mobility, malfunction and so on, the node should try to associate with new parent in its POS [11]. Meanwhile if the current node realizes that the next-hop node regardless whether its child or parent node is not accessible for some reason, the node shall try to recover the path or to report this forwarding error to the source of the packet.

Even though a route maintenance mechanism has been defined in HiLOW, the mechanism is seen as not sufficient to maintain the routing tree. An Extended Hierarchical Routing Over 6LoWPAN which extends HiLOW was presented by in [16] in order to have better maintained routing tree. They suggested two additional fields to be added to the existing routing table of HiLOW namely, Neighbour_Replace_Parent (NRP) and Neighbour_Added_Child (NAC). This NRP doesn't point to the current parent node but to another node which can be its parent if association to current parent fails. Meanwhile NAC refers to the newly added child node. More work need to be done on this mechanism on how many nodes allowed to be adapted by a parent node in addition to the defined MC and whether this mechanism will have any impact on the routing operation, however this topic is beyond the scope of this paper.

III. DISADVANTEGOUS CHILD NODE ATTACHMENT AVOIDANCE MECHANISM

A disadvantageous child parent attachment avoidance mechanism for HiLOW is being suggested in this paper. The suggested mechanism is able to overcome the bias child node phenomena that could shorten the life span of the network as well as affect the reliability and the stability of the network. We are suggesting a mechanism where the new child node is provided with three data; one is the Link Quality Indicator (LQI) value, secondly the depth of the potential parent node and thirdly the average amount of energy the potential parent node has. The suggested mechanism is an enhancement work of the mechanism suggested in [17].

LQI value can be measured by either parent node or by child node itself. The prior measurement by potential parent node was more preferred compared to the latter which is measurement by child. The measurement by child was not selected as the child node needs to use energy to measure LQI for every potential parent node and this would make the total energy usage to be higher compared to energy used when parent node measures the LQI only for one time.

Therefore in this mechanism the LQI value will be measured by the potential parent node and the value is provided to the new node which is looking for a parent node; this is in contrast to the mechanism suggested in [18]. LQI is selected compared to Received Signal Strength Indication (RSSI) as LQI is more accurate to measure the quality of the link and the delivery ratio especially when obstructions or noise exist [18]. In previous mechanisms quality of the link is not considered in selecting the parent node. The quality of the link is important in making association as bad links would cause retransmission of data to occur and this causes nodes to use more energy.

The node which is able to accept the new node as child node will calculate the average amount of energy it has according to the mathematical equation in (4) which is defined earlier in [17]. The average amount of energy represents the energy the potential parent node can equally use for itself as well as use to forward the data of existing child nodes and potential child. So the value 2 in the equation represents itself and the potential new child node.

Avg : Average Amount of Energy

CBP : Current Energy Level of the Potential Parent

EC : Existing Child Node

$$Avg = CBP / (EC+2) \quad (4)$$

In situation where there is more than one potential parent the child node will then make decision on which potential parent node to associate with according to steps as displayed in Fig. 1. First the child node will calculate the average LQI based on the equation (5). The '[]' symbol represents floor operation The average is calculated by summing up all the LQI value received, then dividing it by the number of parent node which responded and lastly flooring the value . In [18] a threshold was to be set and the LQI's will be compared to this threshold, this method needs intervention from human in determining the threshold and a method to communicate the threshold to all nodes; the communication process will again consume energy, due to this factor the new method is introduced.

ALQI : Average LQI of all the potential links

TLQI : Total LQI of all the potential links

Count : Count of potential parent node

$$ALQI = [TLQI / Count] \quad (5)$$

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

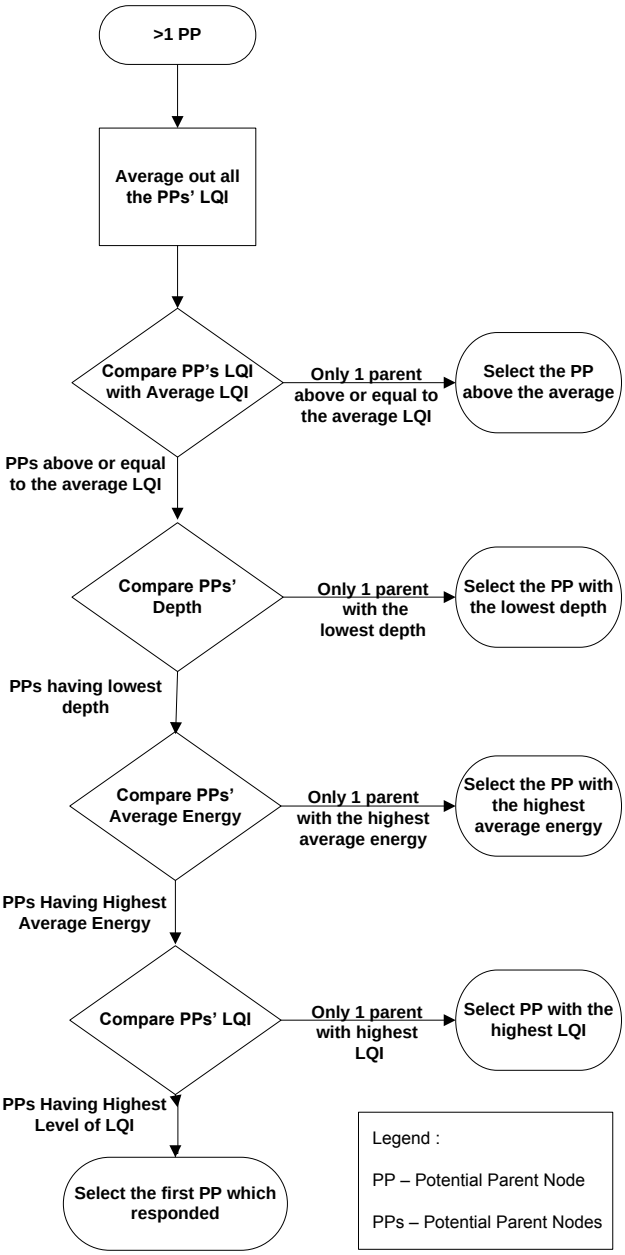


Figure 1. Parent Node Selection Mechanism

The child node then compares all the potential parents' LQI with the average LQI calculated. In the case only one parent node's LQI is higher than average LQI then it would be selected as parent node. If there is more than one parent above the average LQI, then the child node will compare the depth of all the potential parent nodes which is above the average. If there is only parent node with the lowest depth, then that particular parent node is selected and process of associating is started. In the case where there is more than one potential parent with the lowest level of depth, then the average energy of these nodes are compared. The child selects the parent with the highest average energy and associate with it if there is only one particular potential parent with highest energy. In a situation where more than one parent shares the highest level of energy and lowest level of depth; the child node will then compare the LQI value of these nodes. The node with highest LQI value will be selected as parent node to be associated with if only one node qualifies. If more than one parent has the lowest level of depth, highest level of energy and highest level of LQI then the child node will try to establish association with the first responded potential parent in this category. In [12] the LQI value is not considered when more than one parent shares the highest level of energy and lowest level of depth, the child straight away selects the first responded parents which is not advantageous if the first responded parent has lower LQI compared to the other potential parent.

The first scenario will be described with the assistance of Fig. 2. When three potential parent node(8), node(4) and node(3) responds; according to the previous mechanism the child node should attach to parent node(3) as it has no child node compared to node(4) which has 2 child node and node(8) which has 1 child node. According to our mechanism the child will not consider the node(3) for attachment as the link quality is below the average, this represent the quality is bad compared to node(4) and node(8) and this could lead to high retransmission rate compared to node(4) and node(8) link. If an assumption that the node(3) has also two child nodes is made, then according to the previous mechanism the node(8) will be selected as suitable parent and this is also disadvantageous. Meanwhile according to the suggested mechanism in both scenarios where the node(3) has none or two child nodes, node(4) will still be selected as the parent node as it above the average LQI and it only goes through 1 hop to sink node compared to 2 hops if attached to node(8). By attaching to node(4), only node(X)'s and node(4)'s energy will be used in transmitting the data to the sink node, meanwhile if attachment in made to node(8) then energy of node(X)'s, node(8)'s and node(1)'s will be consumed in transmitting the data to the sink node.

The second scenario is when there are two or more potential parent nodes with different number of existing child nodes as represented in Fig. 4. According to the previous mechanism the node(X) should associate itself with node(17) or node(3) based on which node responded first as both nodes have no child node compared to node(8) which has one child node. In the case the first responding parent is node(3) then it is disadvantageous if the child join node(3) as the link quality is not good compared to node(17). Meanwhile if the first responded parent node is node(17) the child will choose node(17) to attach with. This attachment is acceptable if all the nodes have same level of energy. In the case node(8) has abundant amount of energy compared to node(17) then this association is disadvantageous. Our mechanism suggests the node(X) to take into consideration the average amount of power the potential parent node has. In the case parent node(8) has more average power than parent node(17) then node(X) will join parent node(8). Meanwhile if the parent node(17) has more power then it will attach itself with parent node(17).

The third scenario is when all potential parents have the same number of child nodes. The previous mechanism didn't anticipate such a situation will occur. Following our mechanism node(X) will first compare potential parent depth from sink node in this case node(4), node(8) and node(17) (Fig.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

3). In this case the node(X) will try to associate with node(4) as it nearer to the sink node and will not consider node(3) has LQI below average. Meanwhile if the potential parent node above average which responded is only node(8) and node(17), node(X) will join the node which has highest average energy. In the case both node(8) and node(17) has the same amount of average energy it will associate itself with the node which has highest LQI. In the case node(8) and node(17) has the same depth, same average energy and same level of LQI then node(X) will attach to the first potential parent which responded.

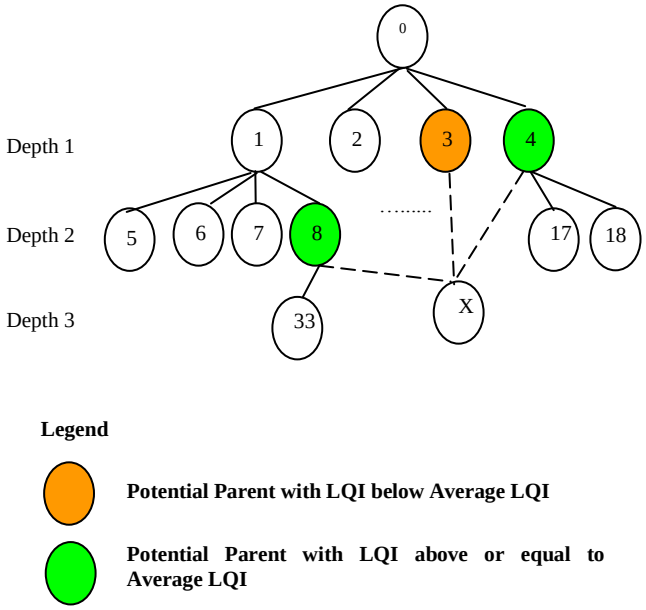


Figure 2. Two Potential Parent Node with LQI above average, with different depth level and different number of existing child node

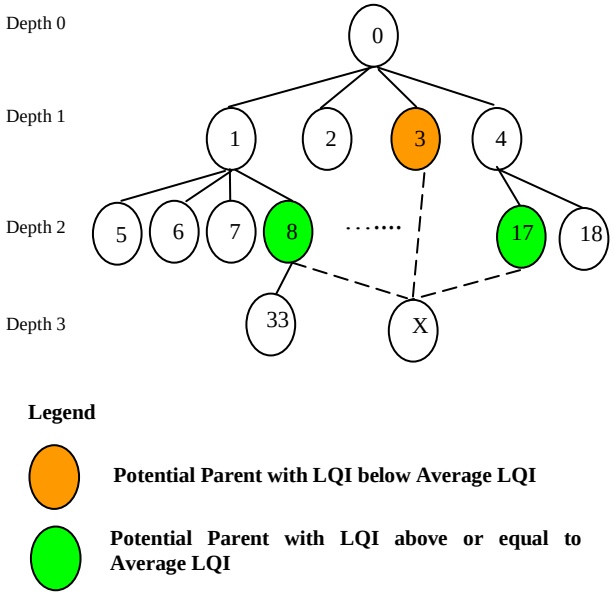


Figure 3. Two potential Parent node with LQI above average and with same depth level but different number of existing child node

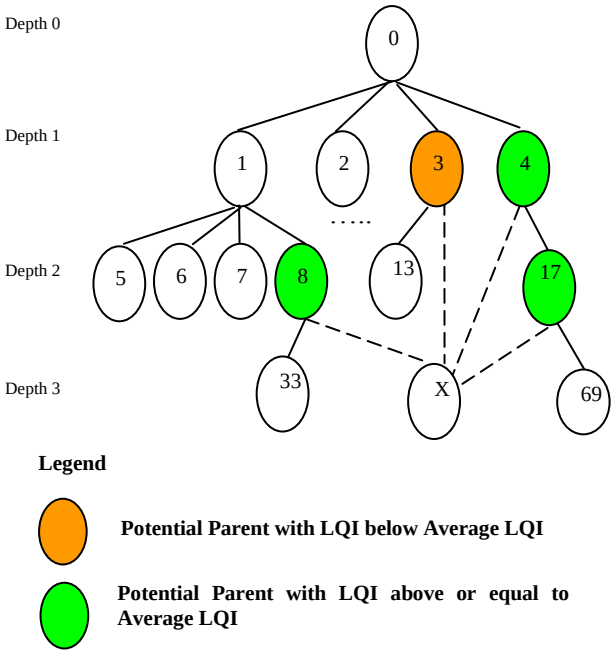


Figure 4. Potential parents with same number of child nodes with some from different depth and some sharing same depth

IV. CONCLUSION

We provides review on HiLOW routing protocol, issues revolving HiLOW and other works done in improving HiLOW. In this paper we have suggested a new mechanism to overcome disadvantageous parent and child node attachment in HiLOW; disadvantageous attachment could jeopardize the reliability of the network, shorten the life span of the network and also cause wastage of energy due to retransmission. Previous mechanism used seems to be less advantageous as it is not considering potential link quality, the potential parent depth and also the energy level. To verify our mechanism, we identify a number of scenarios that need to optimize LQI, average energy level and parent-child nodes parameters. By analyzing these scenarios, the proposed new mechanism is shown to optimize the parameters better that leads to establishing of a more reliable network and enhancing the lifetime of the network. However, the implementation of the proposed mechanism in an actual network is a future work. Our future research will be focused on validating the suggested mechanism as well as solving the other issues highlighted in this paper.

ACKNOWLEDGMENT

The author would like to acknowledge Universiti Sains Malaysia (USM) for funding of USM Fellowship Scheme 2009/10.

REFERENCES

[1] Ian F.Akyildiz, Weilian Su, Yogesh Sankarasubramaniam and Erdal Cayirci, “A Survey on Sensor Networks”, Communication Magazine, IEEE, Volume 40

[2] A. Dunkels, F. ¨Osterlind, N. Tsiftes, and Z. He. Software-based online energy estimation for sensor nodes. In Proceedings of the Fourth IEEE

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Workshop on Embedded Networked Sensors (Emnets IV), Cork, Ireland, June 2007.

[3] L.Martin, B.D Mads, B.Philippe "Bluetooth and sensor networks: a reality check", Proceedings of the 1st international conference on Embedded networked sensor systems, 2003

[4] S.Jianping, et al., "WirelessHART: Applying Wireless Technology in Real-Time Industrial Process Control", In Proceedings of IEEE Real-Time and Embedded Technology and Applications Symposium, 2008.

[5] B.Chiara, C.Andrea, D.Davide, V.Roberto, "An Overview onWireless Sensor Networks Technology and Evolution", Sensors 2009, Sensors 2009, 9, 6869-6896; doi:10.3390/s90906869

[6] N. Kushalnagar, et al., "Transmission of IPv6 Packets over IEEE 802.15.4 Networks", rfc4944, September 2007.

[7] IEEE Computer Society, "802.15.4-2006 IEEE Standard for Information Technology- Telecommunications and Information Exchange Between Systems- Local and Metropolitan Area Networks- Specific Requirements Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs)"

[8] K. Kim, S.Yoo, S.Daniel, J.Lee, G.Mulligan, "Problem Statement and Requirements for 6LoWPAN Routing", draft-ietf-6lowpan-routing-requirements-02, March 2009

[9] K. Kim, S.Yoo, S.Daniel, J.Lee, G.Mulligan, "Commisioning in 6LoWPAN", draft-6lowpan-commisioning-02, July 2008

[10] K. Kim, et al., "Hierarchical Routing over 6LoWPAN (HiLOW)", draft-daniel-6lowpan-hilow-hierarchical-routing-01, June 2007.

[11] K. Kim, et al., "Hierarchical Routing over 6LoWPAN (HiLOW)", draft-daniel-6lowpan-hilow-hierarchical-routing-00, June 2005.

[12] K. Kim, G.Montenegro, S.Park, I.Chakeres, C.Perkins, "Dynamic MANET On-demand for 6LoWPAN (DYMO-low) Routing", draft-montenegro-6lowpan-dymo-low-routing-03, June 2007.

[13] K.Kim, S.Daniel, G.Montenegro, S.Yoo, N.Kushalnagar, "6LoWPAN Ad Hoc On-Demand Distance Vector Routing (LOAD", draft-daniel-6lowpan-load-adhoc-routing-02, March 2006

[14] Martin Haenggi, "Opportunities and Challenges in Wireless Sensor Network", Sensor Network Protocol, Taylor & Francais Group pp. 1-1,1-7.

[15] Hun-Jung-Lim, Tai-Myoung Chung, "The Bias Routing Tree Avoiding Technique for Hierarchical Routing Protocol over 6LoWPAN", 2009 Fifth International Joint Conference on INC, IMS and IDC.

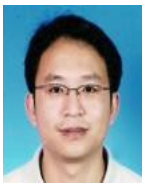
[16] C.Nam, H.Jeong, D.Shin, "Extended Hierarchical Routing Protocol over 6LowPAN", MCM2008, September 2008.

[17] V.C.Lingeswari et al., "Bias Child Node Association Avoidance Mechanism for Hierarchical Routing Protocol in 6LoWPAN", Proceedings of the Third IEEE International Conference on Computer Science and Information Technology (In Press)

[18] Zhu Jian, Zhao Lai, "A Link Quality Evaluation Model in Wireless Sensor Networks", Proceedings of the 2009 Third International Conference on Sensor Technologies and Applications.

AUTHORS PROFILE

Lingeswari V.Chandra was born in Penang, Malaysia. She obtained her BIT with Management degree from AIMST University in 2008. She is the university gold medalist. She obtained her software engineering foundation training from Infosys, Bangolore. She is currently pursuing her PhD in National Advanced IPv6 Center, Universiti Sains Malaysia.



Kok-Soon Chai was born in Penang, Malaysia. He is a certified Project Management Professional, Project Management Institute, USA. He received his MSc and Ph.D. (2003) degrees from the University of Warwick, UK. He worked for more than seven years as a senior R&D software engineer, embedded software manager, and CTO at Motorola, Agilent, Plexus Corp., Wind River in Singapore (now a division of Intel Corp.), and NeoMeridian. He holds one US patent, with two US patents pending. His main interests are wired and wireless sensor networks, green technology, embedded systems, consumer electronics, and real-time operating systems. Dr. Chai is a senior lecturer at the National Advanced IPv6 Centre of Excellence (NAV6) in Universiti Sains Malaysia.

Gopinath Rao Sinniah obtained his BComp Science and MSc (Comp Science) from the Universiti Sains Malaysia in 1999 and 2004 and currently pursuing his Ph.D at the same university. He has involved in the IPv6 development work since 1999 and currently working at MIMOS Berhad as Senior Staff Researcher focusing on IPv6 specifically on wireless sensors network. Currently he holds 10 patents filed locally and 1 at WIPO. He is also the chairman of MTSFB IPv6 working group.

Sureswaran Ramadass obtained his BsEE/CE (Magna Cum Laude) and Master's in Electrical and Computer Engineering from the University of Miami in 1987 and 1990, respectively. He obtained his Ph.D. from Universiti Sains Malaysia (USM) in 2000 while serving as a full-time faculty in the School of Computer Sciences. Dr. Sureswaran Ramadass is a Professor and the Director of the National Advanced IPv6 Centre of Excellence (NAV6) in Universiti Sains Malaysia.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Rough Entropy as Global Criterion for Multiple DNA Sequence Alignment

Sara El-Sayed El-Metwally
Dept. of Computer Science
Faculty of Computer and Information Science
Mansoura, Egypt
Sarah_almetwally4@yahoo.com

ElSayed Radwan
Dept. of Computer Science
Faculty of Computer and Information Science
Mansoura, Egypt
elsfradwan@yahoo.com

Taher. Hamza
Dept. of Computer Science
Faculty of Computer and Information Science
Mansoura, Egypt
Taher_hamza@yahoo.com

***Abstract*-This paper presents a new method for multiple sequence alignment using rough entropy as a global criterion to measure the quality of alignment. This method collects *DNA* sequences in clusters based on rough sets indcernibility relation. Rough entropy is used to maximize the total number of sequences inside each cluster with respect to the total number of sequences being aligned. The method terminates when all aligned sequences are located in all clusters and hence the problem is near optimally solved.**

Keywords-Multiple Sequence Alignment MSA, Rough Set Theory, Indcernibility relation,Rough Entropy, Similarity relation SIM(P),$S_P(x)$.

I. INTRODUCTION

Multiple Sequence Alignment (*MSA*) is a fundamental and challenging problem in computational molecular biology. It plays a key role in computing similarity and in finding highly conserved subsequences among set of *DNA* sequences .*MSA* is one of the most commonly used methods for inferring biological structures and functions. Moreover, it is the first step of many tasks in computational biology involving fragment assembly, evolutionary tree reconstruction, and genome analysis.

For pairwise alignment, algorithms such as Needleman-Wunsch and Smith-Waterman [1] which use dynamic programming approach, guarantee finding the optimal solution [2] .For multiple alignments, methods such as multidimensional and repeating pairwise dynamic programming, try to find near optimal solution in reasonable amount of time. Finding optimal solution in reasonable amount of time becomes difficult when the number of sequences and the length of each sequence increase. The problem of computing minimum cost for *MSA* has been shown to *NP*-hard. As a result, all multiple alignment algorithms currently in use depend on different kinds of heuristics [2, 4, 7, and 8].

The most common current solution to multiple *DNA* sequence alignment problem is to use an evolutionary computation (*EC*) where genetic algorithms find better alignment of sequences with better similarity score by Darwin's theory about evolution. Genetic algorithms use an objective function to evaluate the quality of alignment. The objective function summarizes the biological knowledge that is intended to be projected into the alignment.

An alignment is considered to be correct if it reflects evolutionary history of the species of sequences being aligned. But, at the time of assessing the quality of alignment, such evolutionary information is not frequently available or even more not known [3].

One of existing measures for assessing the quality of alignment is obtained by computing similarity using substitutions matrices (for proteins). A substitution matrix assigns a cost for each possible substitution or conservation accordingly to the probability of occurrence computed from data analysis. In this approach insertions and deletions are weighted using *affine gap penalties* model. This model assigns weight cost for each gap opening and gap extension in order to favor alignments with smaller numbers of indels (each gap can be regarded as an insertion-deletion event). The main disadvantage of these substitution matrices is that they are intended to rate the similarity between two sequences at a time only. In order to extend them to multiple sequences, they are scaled by adding up each pairwise similarity to obtain the score of multiple sequence alignment [3, 4]. Another existing measure is to use *weighted sum of pairs*, is the objective function used by *MSAs* that associates a cost for each pair of aligned bases in each column of the alignment (substitution cost) and another similar cost for gaps (gap cost), sum of these costs yields a global cost of the alignment[8,7 and 11].

This paper presents a new approach for multiple sequence alignment using Rough Entropy as Global Criterion, *REGC*, to measure the quality of alignment. This approach doesn't use any traditional methods that compute similarity between all sequences using scoring functions, Rather than rough sets indcernibility relation used to collect DNA sequences in clusters .Each cluster contains set of sequences that are similar to each other. Gaps are inserted to each DNA sequence according to specified length then sequences are aligned

randomly by change offsets of gaps in each sequence. In each epoch, the random alignment is converted to suitable representation for rough set analysis, matrix form, where each sequence represents one row and each position of nucleotide inside each sequence represents one column. The value of each cell in the matrix can be a nucleotide base or gap according to random alignment. Indcernibility definition for missing values is used to compute the maximum clustering of sequences due to the presence of gaps. *REGC* used to measure the quality of alignment by maximize total number of sequences inside each cluster with respect to the total number of sequences. Each cluster has a name of one sequence and objects inside it are set of sequences that are indiscernible by knowledge P , set of attribute-value pairs which are corresponding to positions and values of bases inside each sequence.

This paper introduces a global criterion to measure the quality of alignment by filling each cluster with the maximum number of sequences in alignment problem. The rough entropy value of knowledge P increases if the total number of sequences inside each clusters increases. When one cluster doesn't contain the maximum number of sequences, rough entropy approach finds similar missing sequences by dropping attributes from the knowledge P . If entropy value of knowledge P after dropping attributes increases, this means that the columns corresponding to those attributes must be realigned. The process repeat until the entropy value remains constant and hence the maximum numbers of sequences that are similar to each other are located inside each cluster.

Paper is organized as follows: Section II gives an overview about preliminaries of *DNA* sequence alignment and rough sets. Section III presents the constructed hybrid model to align *DNA* sequences based on Rough sets and Rough Entropy. Section IV show simple experimented results of proposed model and section V conclude this paper.

II. PRELIMINARIES

A. DNA Sequence Alignment

DNA consists of two long strands of simple units called nucleotides, with backbones made of sugars and phosphate groups joined by ester bonds. These two strands run in opposite directions to each other and are therefore anti-parallel. Attached to each sugar is one of four types of letters called bases. The four bases found in *DNA* are adenine (abbreviated *A*), cytosine (*C*), guanine (*G*) and thymine (*T*). Each type of base on one strand forms a bond with just one type of base on the other strand. This is called complementary base pairing, with *A* bonding only to *T*, and *C* bonding only to *G*. The sequence of these four bases carries out the genetic information that is used in the development and functioning of all known living organisms, Thus *DNA* is the long-term storage of genetic information. *DNA* is copied and transmitted from parent to child, but from one generation to the next errors in copying can occur, with one nucleotide being replaced by another, or a new nucleotide being inserted, or an existing nucleotide being deleted [10,12 and 13].Over many generations, organisms with a common ancestor can end up with fairly different *DNA* sequences.

To align two or more sequences, they are put together in a $S \cdot C$ matrix, where *S* is the number of sequences, and *C* is the maximum number of bases in a sequence (positions in the alignment); shorter sequences are filled at the end with gap ("–") to fit the matrix perfectly. The goal of alignment process is to try to line up as many letters or portions of sequences that are the same, to minimize the number of substitutions, insertions, and deletions that would have had to happen if the sequences came from a common ancestor. In order to choose best alignment, the process of alignment has an objective to align homologous residues (having the same evolutionary origin) .The number of possible alignments of two sequences grows exponentially as the length of sequences increases [3].So, with more sequences involved in the alignment process, the number of possible alignments grows faster and the problem of find optimal alignment becomes difficult to solve [3].

B. Rough Set Theory

Rough set theory proposed by Pawlak [9] is an effective approach to imprecision, vagueness, and uncertainty. Rough set theory overlaps with many other theories such that fuzzy sets, evidence theory, and statistics. From a practical point of view, it is a good tool for data analysis. The main goal of the rough set analysis is to synthesize approximation of concepts from acquired data. The starting point of rough set theory is an observation that objects having the same description are indiscernible (similar) with respect to the available information. Determination of similar objects with respect to the defined attributes values is very hard and sensible when some attribute values are missing. This problem must be handled very carefully. The indiscernibility relation is a fundamental concept of rough set theory which used in the complete information systems. In order to process incomplete information systems, the indiscernibility relation needs to be extended to some equivalent relations.

The starting point of rough set theory which is based on data analysis is a data set called information system (IS). IS is a data table, whose columns are labeled by attributes, rows are labeled by objects or cases, and the entire of the table are the attribute values. Formally, $IS = (U, AT)$, where U and AT are nonempty finite sets called "the universe" and "the set of attributes," respectively. Every attribute $a \in AT$, has a set of V_a of its values called the "domain of a ". If V_a contains missing values for at least one attribute, then S is called an incomplete information system, otherwise it is complete [5, 6, 8].

Any information table defines a function ρ that maps the direct product $U \times AT$ into the set of all values assigned to each attribute. The example of incomplete information system depicted in ***Table1*** where set of objects in the universe corresponding to set of DNA sequences and set of attributes corresponding to set of bases inside each sequence. The values

of attributes are corresponding to the values of bases inside each sequence such as the value of Sequence1 at Base0 is defined by $\rho(\text{Seq}_1, \text{Base}_0) = \text{G}$

TABLE 1: Example of Incomplete Information System

	Base_0	Base_1	Base_2
Seq_1	G	A	-
Seq_2	G	T	A
Seq_3	C	C	A

The concept of the indiscernibility relation is an essential concept in rough set theory which is used to distinguish objects described by a set of attributes in complete information systems. Each subset A of AT defines an indiscernibility relation as follows:

$$IND(A) = \{(x, y) \in U \times U : \rho(x, a) = \rho(y, a) \quad \forall a \in A, A \subset AT\} \tag{1}$$

The family of all equivalence classes of $IND(A)$ is denoted by $U / IND(A)$ or U / A [5, 6 and 8].

Obviously $IND(A)$ is an equivalence relation and:

$$IND(A) = \bigcap IND(a) \ \ where\ a \in A \tag{2}$$

A fundamental problem discussed in rough sets is whether the whole knowledge extracted from data sets is always necessary to classify objects in the universe; this problem arises in many practical applications and will be referred to as knowledge reduction. The two fundamental concepts used in knowledge reduction are the core and reduct. Intuitively, a reduct of knowledge is essential part, which suffices to define all basic classifications occurring in the considered knowledge, whereas core is in a certain sense the most important part. Let A set of attributes and let $a \in A$, the attribute a is dispensable in A if:

$$IND(A) = IND(A - \{a\}) \tag{3}$$

Otherwise a is indispensable attribute. The set of attributes B, where $B \subset A$ is called reduct of A if:

$$IND(B) = IND(A) \tag{4}$$

A may have many reducts. The set of all indispensable attributes in A will be called the core of A, and will be denoted as $CORE(A)$:

$$CORE(A) = \bigcap RED(A) \tag{5}$$

Recently A.Skowron [8] has proposed to represent knowledge in a form of discernibility matrix .this representation has many advantages because it enables simple computation of the core and reduct of knowledge.

Let $K = (U, A)$ be a knowledge representation system with $U = \{x_1, x_2, \ldots, x_n\}$ by a discernibility matrix of K denoted by $M(k)$, which means $n \times n$ matrix defined by:

$$(C_{ij}) = \{a \in A : \rho(x, a) \neq \rho(y, a)\} \ for\ i, j = 1,2....,n. \tag{6}$$

Thus entry c_{ij} is the set of all attributes which discern objects xi and xj.

The core can be defined now as the set of all single element entries of the discernibility matrix, i.e.

$$CORE(A) = \{a \in A : C_{ij} = (a)\ for\ some\ i,j\}. \tag{7}$$

It can be easily seen that $B \subset A$ is the reduct of A if B is the minimal subset of A such that $B \bigcap c \neq \phi$ for any nonempty entry $c\ (c \neq \phi)$ in $M(k)$.In other words reduct is the minimal subset of attributes that discerns all objects discernible by the whole set of attributes.Let $C, D \subset A$ be two subsets of attributes, called condition and decision attributes respectively. KR-system with distinguished condition and decision attributes will be called a decision table and will be denoted $T = (U, A, C, D)$. Every $x \in U$ associate a function $d_X : A \to V_a$, such that $d_X(a) = a(x)$, for every $a \in C \bigcup D$; the function d_X will be called a decision rule, and x will be referred to as a label of the decision rule d_X [5, 6, and 9].

III. Hybrid Model of Rough Sets and Rough Entropy

The hybrid model discussed in this paper is considered to be a combination system that contains two methodologies which are rough sets and rough entropy. *DNA* sequences are converted to suitable representation for rough set analysis. The rough set indcernibility relation used to collect *DNA* sequences in clusters .Each cluster contains set of sequences that are similar to each other. Rough Entropy is used to measure the quality of alignment by maximize the total number of sequences inside each cluster with respect to the total number of sequences being aligned.

A. Rough Sets analysis for Sequence Alignment

The first part of the model consists of rough set analysis for *DNA* sequence alignment. Rough set approach used here is modified to deal with incomplete information system, where $IIS = (U, AT)$, where U and AT are nonempty finite sets called "the universe" and "the set of attributes," respectively. Every attribute $a \in AT$, has a set of values called V_a and this set contains missing values for at least one attribute. In order to process incomplete information systems (IIS), the indiscernibility relation has been extended to some equivalent relations such as similarity relation. Similarity relation $SIM(P)$ denotes a binary relation between objects that are possibly indiscernible in terms of values of attributes and in the case of missing values the modified relation is defined by equation **(8):**

Identify applicable sponsor/s here. *(sponsors)*

$$SIM(P) = \{(x,y) \in U \times U,\ a \in P, \rho(x,a) = \rho(y,a) \quad or \quad \rho(x,a) = * \ or \ \rho(y,a) = *\} \quad \textbf{(8)}$$

$$SP(x) = \{y \in U : (x,y) \in SIM(P), P \subset AT\} \quad \textbf{(9)}$$

$SP(x)$ Denotes maximal set of objects which are possibly indiscernible by P with x [5, 6]. The indcernibility relation in rough sets, as depicted in equation ***(8),*** used to collect DNA sequences in clusters .Each cluster contains set of sequences that are similar to each other. Gaps are inserted to each DNA sequence according to specified length as depicted in ***Figure 1*** then the sequences are aligned randomly by change offsets of gaps in each sequence as depicted in ***Figure 2***.

```
G C A T G C T A - - - - - - - - - - - -
A G C T G C - - - - - - - - - - - - - -
T A G C A A - - - - - - - - - - - - - -
G C A CA T T - - - - - - - - - - - - -
```

Figure 1: Gap insertion for Length=20 bases

```
---GC--A-TG--C---T-A
A--G----CTG--C------
-TAGC--A---A--------
--G-C-A--CATT-------
```

Figure 2: Random Alignment of DNA Sequences

In each epoch, the random alignment is converted to suitable representation for rough set analysis, matrix form, where each sequence represents one row and each position of nucleotide inside each sequence represents one column. The value of each cell in the matrix can be a nucleotide base or gap according to random alignment as depicted in ***Table 2***.

TABLE 2: Snapshot of Rough Representation Table

(a)

U	P_0	P_1	P_2	P_3	P_4
S_1	*	*	*	G	C
S_2	A	*	*	G	*
S_3	*	T	A	G	C
S_4	*	*	G	*	C

(b)

U	P_5	P_6	P_7	P_8	P_9
S_1	*	*	A	*	T
S_2	*	*	*	C	T
S_3	*	*	A	*	*
S_4	*	A	*	*	C

Indcernibility definition of missing values is used here to compute the maximum clustering of sequences due to the presence of gaps as depicted in ***Figure 3***.

```
the Similar Sequences for Se0=  0 1 2
the Similar Sequences for Se1=  0 1 2
the Similar Sequences for Se2=  0 1 2
the Similar Sequences for Se3=  3
```

Figure 3: Set of Clusters corresponding to one Epoch

B. Rough Entropy for assessing the Quality of Alignment

The second part of the model uses ***REGC*** to measure the quality of *DNA* sequence alignment. The definition of rough entropy of knowledge in incomplete information system has been introduced as:

Let $IIS = (U, AT)$, $P \subset AT$ the rough entropy of knowledge P is defined by the following equation:

$$E(P) = -\sum_{i=1}^{|U|} \frac{|S_P(x_i)|}{|U|} \log \frac{1}{|S_P(x_i)|} \quad \textbf{(10)}$$

Where:

- $U = \{x_1, x_2, x_3, ..., x_U\}$, set of objects in the universe.
- $|U|$ is the cardinality of set U .
- $Log\ x = log_2\ x$.
- $\frac{|S_P(x_i)|}{|U|}$ represents the probability of tolerance class $S_P(x_i)$ within the universe U .
- $\frac{1}{|S_P(x_i)|}$ denotes the probability of one of values in tolerance class $S_P(x_i)$.

The maximum value of rough entropy for knowledge P is computed by equation **(11)**:

$$E(P) = |U| \log |U| \quad \textbf{(11)}$$

This value is achieved only by the equation **(12)**:

$$U/SIM(P) = \{S_P(x) = U \mid x \in U\} \quad \textbf{(12)}$$

The minimum of rough entropy for knowledge P is 0 . This value is achieved only by the equation **(13)**:

$$U/SIM(P) = \{S_P(x) = \{x\} \mid x \in U\}\ [6]. \quad \textbf{(13)}$$

REGC is used to measure the quality of alignment by maximize the total number of sequences inside each cluster with respect to the total number of sequences. Each cluster has a name of one sequence such as depicted in ***Figure 3*** and the objects inside it are set of sequences that are indiscernible by knowledge P , set of attribute-value pairs which are corresponding to positions and values of bases inside each sequence, such as depicted in ***Table 2***.

The aim of our approach is to make a global criterion to measure the quality of alignment by filling each cluster with

the maximum number of sequences in alignment problem as depicted in ***Figure 4*** where the total number of sequences being aligned is 4 sequences and the total number of clusters is 4 clusters one for each sequence .***REGC*** try to fill each cluster with maximum number of sequences similar to the sequence corresponding to that cluster. The alignment process ends when the total number of sequences located inside each cluster is equal to the total number of sequences being aligned.

```
the Similar Sequences for Se0=  0 1 2 3
the Similar Sequences for Se1=  0 1 2 3
the Similar Sequences for Se2=  0 1 2 3
the Similar Sequences for Se3=  0 1 2 3
```

Figure 4: Maximum Clusters Set

When the total number of sequences inside each cluster increases, the rough entropy value of knowledge P increases such as depicted in ***Figure 5*** where the rough entropy value of P in ***Figure 5-b*** is higher than in ***5- a*** because clusters in ***5-b*** are filling with the maximum number of sequences that are similar to each other.

```
the Similar Sequences for Se0=  0 1 2
the Similar Sequences for Se1=  0 1 2
the Similar Sequences for Se2=  0 1 2
the Similar Sequences for Se3=  3
E(A)=3.5661656266226
```

(a)

```
the Similar Sequences for Se0=  0 1 2 3
the Similar Sequences for Se1=  0 1 2 3
the Similar Sequences for Se2=  0 1 2 3
the Similar Sequences for Se3=  0 1 2 3
E(A)=8
```

(b)

Figure 5: Rough Entropy Value for Knowledge P

When one cluster doesn't contain the maximum number of sequences, rough entropy approach finds similar missing sequences by dropping attributes from the knowledge P. If the entropy value of knowledge after dropping attributes increases, this means that the columns corresponding to those attributes must be realigned. The process repeat until the entropy value remains constant and hence the maximum number of sequences that are similar to each other are located inside each cluster.

REGC Algorithm for MSA

- *Input:*
  N : Total number of sequences to be aligned.
  M : Max Length.
  R : Randomization Times (optional).
  $S_1, S_2, S_3, ..., S_N$: Sequences being aligned.
- *Output:*
  Alignment of DNA Sequences

1. *[Start]* insert gaps to all sequences until all have the same length M .
2. *[Alignment]*
   - If (R =1) create an alignment by randomly change the offset of gaps in each sequence.
   - If (R =2,3,.. R) create an alignment by randomly change the offset of gaps in columns indexed as realigned from attribute analysis step in previous epoch
3. *[Rough representation]* convert the resulting random alignment to a suitable representation for rough set analysis , matrix form, where each sequence represent one row and each position of nucleotide inside each sequence represent one column. The value of each cell in the matrix can be a nucleotide base or gap according to random alignment. Rows corresponding to set of objects in the universe and columns are set of attributes, knowledge P
4. *[Rough Entropy]* compute the entropy value of knowledge P that satisfies equation 10.
5. *[Attribute Analysis]****REGC*** analysis the entropy value by check each cluster if contains the total number of sequences being aligned and hence for any sequence x $S_P(x) = U$ .When one cluster doesn't contain the maximum number of sequences, ***REGC*** finds similar missing sequences by dropping $1,2,.....,M-1$ attributes from the knowledge P. If the entropy value of knowledge after dropping attributes increases, this means that the columns corresponding to those attributes must be realigned.
6. *[Loop] go to step 2* until the maximum value of rough entropy of knowledge P is $|U|\log|U|$ reached, this value is achieved only by $U/SIM(P) = \{S_P(x) = U \mid x \in U\}$, or until the maximum number of randomization times encountered.

IV. Experimented Results

All *DNA* fragments supplied to our model are obtained from Gene repository over the internet i.e. http://www.ncbi.nlm.nih.gov/Genbank/ [14] Genbank, is the NIH genetic sequence database, an annotated collection of all

publicly available *DNA* sequences. There are approximately 106,533,156,756 bases in 108,431,692 sequence records in the traditional GenBank divisions and 148,165,117,763 bases in 48,443,067 sequence records in the WGS division as of August 2009. GenBank is part of the International Nucleotide Sequence Database Collaboration, which comprises the *DNA* Databank of Japan (DDBJ), the European Molecular Biology Laboratory (EMBL), and GenBank at NCBI. These three organizations exchange data on a daily basis.

Consider these four input fragments of *DNA* sequences which are given from Genbank.

GCATGCTA
AGCTGC
TAGCAA
GCACATT

Where the other input parameters supplied to our program are *N*=4, *M*=20 and *R*=100.

1. *[Start]* insert gaps to all sequences until all have the same length 20.

G C A T G C T A - - - - - - - - - - - -
A G C T G C - - - - - - - - - - - - - -
T A G C A A - - - - - - - - - - - - - -
G C AC A T T - - - - - - - - - - - - -

2. *[Alignment]* create an alignment by randomly change the offset of gaps in each sequence as depicted in ***Figure 6***.

```
---GC--A-TG--C---T-A
A--G----CTG--C------
-TAGC--A---A--------
--G-C-A--CATT-------
the Similar Sequences for Se0=  0 1 2
the Similar Sequences for Se1=  0 1 2
the Similar Sequences for Se2=  0 1 2
the Similar Sequences for Se3=  3
```

Figure 6: Random Alignment of DNA Sequences

3. *[Rough representation]* convert the resulting random alignment to a suitable representation for rough set analysis , matrix form, where each sequence represent one row and each position of nucleotide inside each sequence represent one column. The value of each cell in the matrix can be a nucleotide base or gap according to random alignment. Rows corresponding to set of objects in the universe and columns are set of attributes, knowledge P , as described in ***Table 3:***

TABLE 3: Rough Representation of DNA Sequence Alignment

(a)

U	P_0	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9
S_1	*	*	*	G	C	*	*	A	*	T
S_2	A	*	*	G	*	*	*	*	C	T
S_3	*	T	A	G	C	*	*	A	*	*
S_4	*	*	G	*	C	*	A	*	*	C

(b)

U	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}	P_{15}	P_{16}	P_{17}	P_{18}	P_{19}
S_1	G	*	*	C	*	*	*	T	*	A
S_2	G	*	*	C	*	*	*	*	*	*
S_3	*	A	*	*	*	*	*	*	*	*
S_4	A	T	T	*	*	*	*	*	*	*

4. *[Rough Entropy]* compute the entropy value of knowledge P that satisfies equation 10 as depicted in ***Figure 7***.

```
--G-C-A--CATT-------
---GC--A-TG--C---T-A
A--G----CTG--C------
-TAGC--A---A--------
--G-C-A--CATT-------
the Similar Sequences for Se0=  0 1 2
the Similar Sequences for Se1=  0 1 2
the Similar Sequences for Se2=  0 1 2
the Similar Sequences for Se3=  3
E(A)=3.5661656266226
```

Figure 7: Rough Entropy Value of Knowledge P

5. *[Attribute Analysis]* when one cluster doesn't contain the maximum number of sequences, rough entropy approach finds similar missing sequences by dropping one, two… $M-1$ attributes from knowledge P . Dropping attributes leads to increase the number of sequences inside each cluster and hence the value of entropy increases. The increasing value of entropy after dropping the attributes indicates that some columns in the alignment are needed to realign and $S_P(x) \neq U$ for any sequence x .When dropping one attribute from knowledge P as depicted in ***Figure 8***, the value of entropy remains constant. Every cluster doesn't contain all sequences in the alignment and the equation $S_P(x) \neq U$ satisfied for any sequence x . When dropping the attributes P_2, P_9, P_{10} *and* P_{11} , the value of entropy increased to 8 as depicted in ***Figure 9*** .Increasing the value of entropy indicates that all columns corresponding to those attribute must be realigned to satisfy this value.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

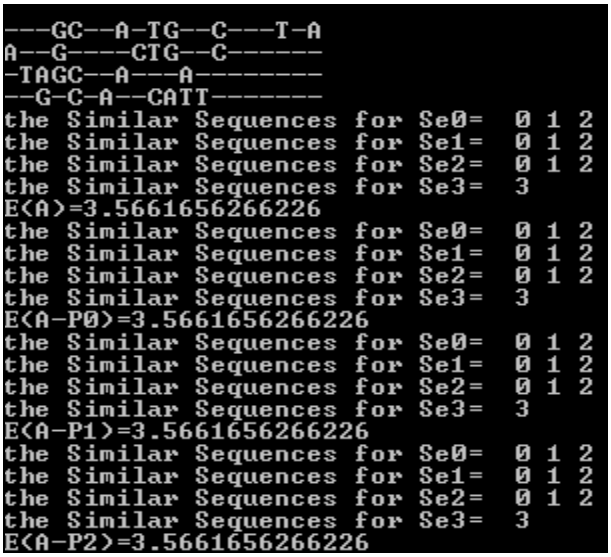


Figure 8: Relation between Dropping Attributes of Knowledge P and Rough Entropy

6. *[Loop] go to step 2* until the maximum value of rough entropy of knowledge P is $|U|\log|U|$ reached as depicted in ***Figure 9*** , this value is achieved only by the $U/SIM(P) = \{S_P(x) = U \mid x \in U\}$, or until the maximum number of randomization times encountered.

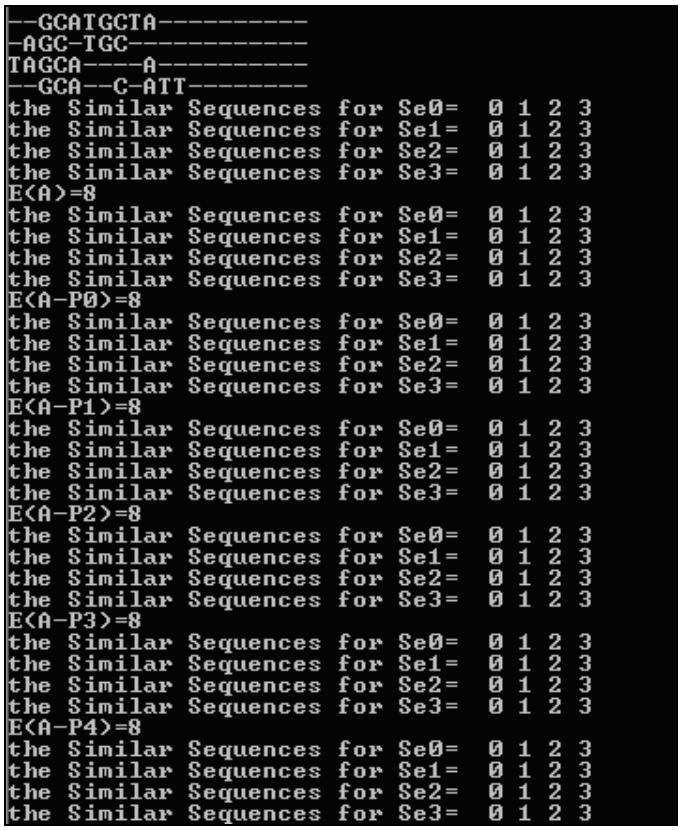


Figure 9: Relation between Dropping Attributes of Knowledge P and Rough Entropy

V. CONCLUSIONS

This paper presents a new approach for multiple sequence alignment using REGC to measure the quality of the alignment. This approach doesn't use any traditional methods that compute similarity between all sequences using scoring functions, Rather than rough set indcernibility relation used to collect DNA sequences in clusters where each cluster contains set of sequences that are similar to each other. Gaps are inserted to each DNA sequence according to specified length then sequences are aligned randomly by change offsets of gaps in each sequence. REGC is used to measure the quality of alignment by maximize the total number of sequences inside each cluster with respect to the total number of sequences being aligned. Each cluster has a name of one sequence and the objects inside it are set of sequences that are indiscernible by knowledge P , set of attribute-value pairs which are corresponding to positions and values of bases inside each sequence. REGC used to measure the quality of alignment by filling each cluster corresponding to one sequence with the maximum number of sequences in the alignment problem .The rough entropy value of knowledge P increases when the total number of sequences inside each cluster increases. If one cluster doesn't contain the maximum number of sequences, rough entropy approach finds similar missing sequences by dropping attributes from knowledge P. If the entropy value of knowledge after dropping attributes increases, this means that the columns corresponding to those attributes must be realigned. The process repeat until the maximum value of rough entropy of knowledge P is $|U|\log|U|$ reached; this value is achieved only by $U/SIM(P) = \{S_P(x) = U \mid x \in U\}$ for any sequence x .

ACKNOWLEDGMENT

I would like to thank my father Dr.El-Sayed El-Metwally and my mother Dr.Hemmat El-Shik for their moral support I required in my life at all.

I am heartily thankful to my supervisor, Dr. ElSayed Radwan, whose encouragement, supervision and support from the preliminary to the concluding level enabled me to develop an understanding of the subject.

Lastly, I offer my regards and blessings to all of those who supported me in any respect during the completion of the paper.

REFERENCES

[1] Arthur M.Lesk, **Introduction to Bioinformatics**, University of Cambridge, USA, Oxford University Press,2002

[2] Bryan Bergeron, **Bioinformatics Computing,** University of British Columbia, USA, Prentice Hall,2002

[3] Edgar D. Arenas-D´ıaz, Helga Ochoterena, and Katya Rodr´ıguez-V´azquez, **Multiple Sequence Alignment Using a Genetic Algorithm and GLOCSA**, Universidad Nacional Aut´onoma de Mexico, Genetic and Evolutionary Computation Conference**,** Proceedings of the 2008 GECCO

conference companion on Genetic and evolutionary computation, pp. 1795–1798, USA, ACM, 2008.

[4] Edward Keedwell and Ajit Narayanan, **Intelligent Bioinformatics**, School of Engineering and Computer Sciences, University of Exeter, UK, Wiley, 2005.

[5] E. A. Rady, M. M. E. Abd El-Monsef, and W. A. Abd El-Latif, **A Modified Rough Set Approach to Incomplete Information Systems,** Journal of Applied Mathematics and Decision Sciences, Article ID 58248**,** Egypt, Hindawi Publishing Corporation, , Volume 2007

[6] JIYE LIAN and ZONGBEN XU, **The Algorithm on Knowledge Reduction in Incomplete Information Systems**, Shanxi University, Taiyuan, International journal of uncertainty, Fuzziness and Knowledge based systems, Volume 2002.

[7] Koji Tajima, **Multiple Sequence Alignment Using Parallel Genetic Algorithms,** Institute for social information Science ,FUJITSU LABORATORIES LTD, Japan ,Genome Informatics Workshop IV, Volume 4, 1993

[8] *Narayanan, E.C. Keedwell and B. Olsson,* **Artificial Intelligence Techniques for Bioinformatics,** School of Engineering and Computer Sciences, University of Exeter, Exeter EX4 4QF, UK, Journal of Applied Bioinformatics , 2005.

[9] ZDZISLAW PAWLAK, **Rough Sets Theoretical Aspects of Reasoning about Data**, Institute of Computer Science, Warsaw University of Technology, Australia, Kluwer Academic Publishers, 1991

[10] http://en.wikipedia.org/wiki/Multiple_sequence_alignment 10-7-2009

[11] http://oreilly.com/news/bioinformatics_0401.html 2-3-2008

[12] http://www.ebi.ac.uk/2can/tutorials/transcription.html 4-4-2008

[13] http://en.wikipedia.org/wiki/Genetic_code 6-5-2008

[14] http://www.ncbi.nlm.nih.gov/Genbank/ 5-1-2010

Weighted Attribute Fusion Model for Face Recognition

S.Sakthivel
Assistant Professor, Department of Information Technology
Sona college of Technology, Salem, India
sakthits@rediffmail.com

Dr.R.Lakshmipathi
Professor, Department of Electrical and Electronic Engineering
St.Peter's Engineering College, Chennai, India
drrlakshmipathi@yahoo.com

***Abstract*—Recognizing a face based on its attributes is an easy task for a human to perform as it is a cognitive process. In recent years, Face Recognition is achieved with different kinds of facial features which were used separately or in a combined manner. Currently, Feature fusion methods and parallel methods are the facial features used and performed by integrating multiple feature sets at different levels. However, this integration and the combinational methods do not guarantee better result. Hence to achieve better results, the feature fusion model with multiple weighted facial attribute set is selected. For this feature model, face images from predefined data set has been taken from Olivetti Research Laboratory (ORL) and applied on different methods like Principal Component Analysis (PCA) based Eigen feature extraction technique, Discrete Cosine Transformation (DCT) based feature extraction technique, Histogram Based Feature Extraction technique and Simple Intensity based features. The extracted feature set obtained from these methods were compared and tested for accuracy. In this work we have developed a model which will use the above set of feature extraction techniques with different levels of weights to attain better accuracy. The results show that the selection of optimum weight for a particular feature will lead to improvement in recognition rate.**

Keywords- Face Recognition, Feature Fusion Method, Parallel Method, PCA, DCT, Histogram Matching

I. INTRODUCTION

Face recognition is an important part of today's emerging biometrics and video surveillance markets. Face Recognition can benefit the areas of: Law Enforcement, Airport Security, Access Control, Driver's Licenses & Passports, Homeland Defense, Customs & Immigration and Scene Analysis. Face recognition has been a research area for almost 30 years, with significantly increased research activity since 1990[16] [15]. This has resulted in the development of successful algorithms and the introduction of commercial products. But, the researches and achievements on face recognition are still in its initial stages of development. Although face recognition is still in the research and development phase, several commercial systems are currently available and research organizations are working on the development of more accurate and reliable systems. Using the present technology, it is impossible to completely model human recognition system and reach its performance and accuracy. However, the human brain has its shortcomings in some aspects. The benefits of a computer system would be its capacity to handle large amount of data and ability to do a job in a predefined repeated manner. The observations and findings about human face recognition system will be a good starting point for automatic face attribute.

A. Early Works

Face recognition has gained much attention in the last two decades due to increasing demand in security and law enforcement applications. Face recognition methods can be divided into two major categories, appearance-based method and feature-based method. Appearance-based method is more popular and achieved great success [3].

Appearance-based method uses the holistic features of a 2-D image [3]. Generally face images are captured in very high dimensionality, normally which is more than 1000 pixels. It is very difficult to perform face recognition based on original face image without reducing the dimensionality by extracting the important features. Kirby and Sirovich first used principal component analysis (PCA) to extract the features from face image and used them to represent human face image [16]. PCA seeks for a set of projection vectors which project the image data into a subspace based on the variation in energy. Turk and Pentland introduced the well-known Eigenface method [15]. Eigenface method incorporates PCA and showed promising results. Another well-known method is Fisher face. Fisher face incorporates linear discriminant analysis (LDA) to extract the most discriminant features and to reduce the dimensionality [3]. when it comes to solving problems of pattern classification, LDA based algorithms outperform PCA based ones, since the former optimizes the low dimensional representation of the objects with focus on the most discriminant feature extraction while the latter achieves simply object reconstruction [9][10][11]. Recently there has been a lot of interest in geometrically motivated approaches to data analysis in high dimensional spaces. This case is concerned with data drawn from sampling a probability distribution that has support on or near a sub manifold of Euclidean space [5].

Let us consider a collection of data points of n-dimensional real vectors drawn from an unknown probability distribution. In increasingly many cases of interest in machine learning and data mining, one is confronted with the situation which is very

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

large. However, there might be reason to suspect that the "intrinsic dimensionality" of the data is much lower. This leads one to consider methods of dimensionality reduction [1][2][8] that allow one to represent the data in a lower dimensional space. A great number of dimensionality reduction techniques exist in the literature.

In practical situations, where data is prohibitively large, one is often forced to use linear (or even sub linear) techniques. Consequently, projective maps have been the subject of considerable investigation. Three classical yet popular forms of linear techniques are the methods of PCA [6] [14] [1][2], multidimensional scaling (MDS) [6] [14], and LDA [14] [11]. Each of these is an eigenvector method designed to model linear variability's in high-dimensional data. More recently, frequency domain analysis methods [3] such as discrete Fourier transform (DFT), discrete wavelet transform (DWT) and discrete cosine transform (DCT) have been widely adopted in face recognition. Frequency domain analysis methods transform the image signals from spatial domain to frequency domain and analyse the features in frequency domain. Only limited low-frequency components which contain high energy are selected to represent the image. Unlike PCA, frequency domain analysis methods are data independent [3]. They analyse image independently and do not require training images. Furthermore, fast algorithms are available for the ease of implementation and have high computation efficiency.

In [3] new parallel models for face recognition were presented. Feature fusion is one of the easy and effective ways to improve the performance. Feature fusion method is performed by integrating multiple feature sets at different levels. However, feature fusion method does not guarantee better result [3]. One major issue is feature selection. Feature selection plays a very important role to avoid overlapping features and information redundancy. New parallel model for face recognition utilizes information from frequency and spatial domains, addresses both features and processes in parallel way. It is well- known that image can be analysed in spatial and frequency domains. Both domains describe the image in very different ways. The frequency domain features [3] are extracted using techniques like DCT, DFT and DWT methods respectively. By utilizing these two or more very different features, a better performance is guaranteed.

Feature fusion method suffers from the problem of high dimensionality because of the combined features. It may also contain redundant and noisy data. PCA is applied on the features from frequency and spatial domains to reduce the dimensionality and extract the most discriminant information [3]. It is surprising that until recent study demonstrated that colour information makes contribution and enhances robustness in face recognition[4].

II. MATERIALS AND METHODS

In statistics, dimension reduction is the process of reducing the number of random variables under consideration, and can be divided into feature selection and feature extraction.

A. Extracting Eigen Features F1

In previous work [1] five algorithms are evaluated, namely PCA, Kernel PCA, LDA [9] [10], Locality Preserving Projections (LPP) [1] [8] and Neighbourhood Preserving Embedding (NPE) [1][2] [5] for dimensionality reduction and feature extraction and found that Kernel PCA was the best performer. This work uses PCA based algorithm to show some improvements in its performance due to the use of Multiple Weighted Facial Attribute Sets. The basic idea of PCA [5] is to project the data along the directions of maximal variances so that the reconstruction error can be minimized. Given a set of data points x1... xn, let a be the transformation vector and $yi = \mathbf{a}^T \mathbf{x}_i$. The objective function of PCA is as follows:

$$a_{opt} = \arg\max_a \sum_{i=1}^{n} (y_i - \bar{y})^2 = \arg\max_a \; a^T C_a \text{ -------(1)}$$

In equation $\bar{y} = \frac{1}{n}\sum y_i$ and C is the data covariance the eigen.The basic functions of PCA are the eigenvectors of the data covariance matrix corresponding to the largest eigenvalues.While PCA seeks direction that are efficient for representation.

B. Extracting DCT Features F2

The DCT [2] can be used to Create DCT feature Set of the Face. The Discrete Cosine Transform is a real domain transform which represents the entire image as coefficients of different frequencies of cosines (which are the basis vectors for this transform). The DCT of the image is calculated by taking 8x8 blocks of the image in Figure 1, which is then transformed individually. The 2D DCT of an image gives the result matrix such that top left corner represents lowest frequency coefficient while the bottom right corner is the highest frequency.

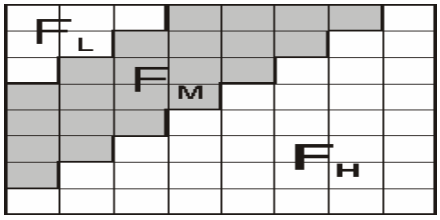


Figure 1 The frequency domain representation of an image

The 1-D *discrete cosine transform* (DCT) is defined as

$$C(u) = \alpha(u) \sum_{x=0}^{N-1} f(x) \cdot \cos\left[\frac{(2x+1)u\pi}{2N}\right] \qquad \text{------- (2)}$$

Similarly, the inverse DCT is defined as

$$f(x) = \sum_{u=0}^{N-1} \alpha(u) C(u) \cdot \cos\left[\frac{(2x+1)u\pi}{2N}\right] \qquad \text{------- (3)}$$

for $x = 0,1,2,\ldots,N-1$. In both **equations (2)** and **(3)** $\alpha(u)$ is defined as

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

$$\alpha(u) = \begin{cases} \sqrt{1/N} & \text{for } u = 0 \\ \sqrt{2/N} & \text{for } u = 1,2,\ldots,N-1 \end{cases} \quad \text{------- (4)}$$

The corresponding 2-D DCT, and the inverse DCT are defined as

$$C(u,v) = \alpha(u)\alpha(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\cdot\cos\left[\frac{(2x+1)u\pi}{2N}\right]\cdot\cos\left[\frac{(2y+1)v\pi}{2N}\right] \text{------ (5)}$$

for $u,v = 0,1,2,\ldots,N-1$ and $\alpha(u)$ and $\alpha(v)$ are defined in **(4)**.

The inverse transform is defined as

$$f(x,y) = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1}\alpha(u)\alpha(v)C(u,v)\cdot\cos\left[\frac{(2x+1)u\pi}{2N}\right]\cdot\cos\left[\frac{(2y+1)v\pi}{2N}\right] \text{-------(6)}$$

The advantage of DCT is that it can be expressed without complex numbers. The DCT transform Equation (5) can be expressed as separable, (like 2-D Fourier transform), i.e. it can be obtained by two subsequent 1-D DCT in the same way as Fourier transform. Equation (6) shows the Inverse transformation.

C. The Histogram Feature Vector F3

The distribution of gray levels occurring in an image is called gray level histogram. It is a graph showing the frequency of occurrence of each gray level in the image versus the gray level itself. The plot of this function provides a global description of the appearance of the image. The histogram of a digital image with gray levels in the range [0, L-1] is a discrete function .

$$P(r_k) = n_k / n \quad \text{------- (7)}$$

Where,

r_k → Kth gray level

n_k → No of pixels in the image with that gray level.

n → total number of pixels in the image.

K = 0, 1, 2… L-1.

L = 256. (For 256 level gray images)

In Equation 7 $P(r_k)$ gives an estimate of the probability of occurrence of gray level r_k .If we use L value of small size, then n_k will contain a range of nearest values in L number f bins. So for constructing Histogram Based Feature, the set n_k and the mid values of the bin n_k were combined.

D. The Intensity Based Feature F4

Intensity Feature set can be formed by using values like mean Median Mode of the 256 gray level images. We can use one or many of this values to represent the average intensity of the face image.

E. Neural Networks and Learning Paradigms

In principle, the popular neural network can be trained to recognize face images directly. However, a simple network can be very complex and difficult to train [12][17]. There are three major learning paradigms, each corresponding to a particular abstract learning task. These are supervised learning, unsupervised learning and reinforcement learning. Usually any given type of network architecture can be employed in any of those tasks.

F. Learning Algorithms

Training a neural network model essentially means selecting one model from the set of allowed models (or, in a Bayesian framework, determining a distribution over the set of allowed models) that minimizes the cost criterion. There are numerous algorithms available for training neural network models; most of them can be viewed as a straightforward application of optimization theory and statistical estimation. Most of the algorithms used in training artificial neural networks are employing some form of gradient descent. This is done by simply taking the derivative of the cost function with respect to the network parameters and then changing those parameters in a gradient-related direction. Evolutionary methods simulated annealing, and Expectation-maximization and non-parametric methods are among other commonly used methods for training neural networks.

G. Support vector machines (SVMs)

Support vector machines are a set of related supervised learning methods used for classification and regression. Viewing input data as two sets of vectors in an n-dimensional space, an SVM will construct a separating hyper plane in that space, one which maximizes the margin between the two data sets[7][13]. To calculate the margin, two parallel hyper planes are constructed, one on each side of the separating hyper plane, which is "pushed up against" the two data sets. Intuitively, a good separation is achieved by the hyper plane that has the largest distance to the neighbouring data points of both classes, since in general the larger the margin the better the generalization error of the classifier. For the linearly separable case, a hyper-plane separating the binary decision classes in the three-attribute case can be represented as the following equation:

$$y = w_0 + w_1x_1 + w_2x_2 + w_3x_3 \quad \text{------ (8)}$$

In Equation (8), y is the outcome x_i, are the attribute values, and there are four weights w_i to be learned by the learning algorithm. In the above equation, the weights w_i are parameters that determine the hyper-plane. The maximum margin hyper-plane can be represented as the following equation in terms of the support vectors:

$$y = b + \sum \alpha_t y_t K(x(t).x) \quad \text{------ (9)}$$

In Equation (9) the function ***K(x(t).x)*** is defined as the kernel function. There are different kernels for generating the inner products to construct machines with different types of nonlinear decision surfaces in the input space.

SVM is selected as the classifying function. One distinctive advantage this type of classifier has over traditional neural networks is that SVMs can achieve better generalization performance. Support vector machine is a pattern classification algorithm developed by Vapnik [13]. It is a binary

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

classification method that finds the optimal linear decision surface based on the concept of structural risk minimization. As shown by Vapnik, this maximal margin decision boundary can achieve optimal worst-case generalization performance. Note that SVMs are originally designed to solve problems where data can be separated by a linear decision boundary

III. THE MODEL OF PROPOSED SYSTEM

Given a set of feature vectors belonging to n classes, a Support Vector Machine (SVM) finds the hyper plane that separates the largest possible fraction of features of the same classes on the corresponding space, while maximizing the distance from either class to the hyper plane. Generally a suitable transformation is first used to extract features of face images and then discrimination functions between each class of images are learned by SVMs. Figure 2 shows the feature set creation for face images from ORL database whereas Figure 3 shows the architecture for training and testing the SVM with weighted attribute set.

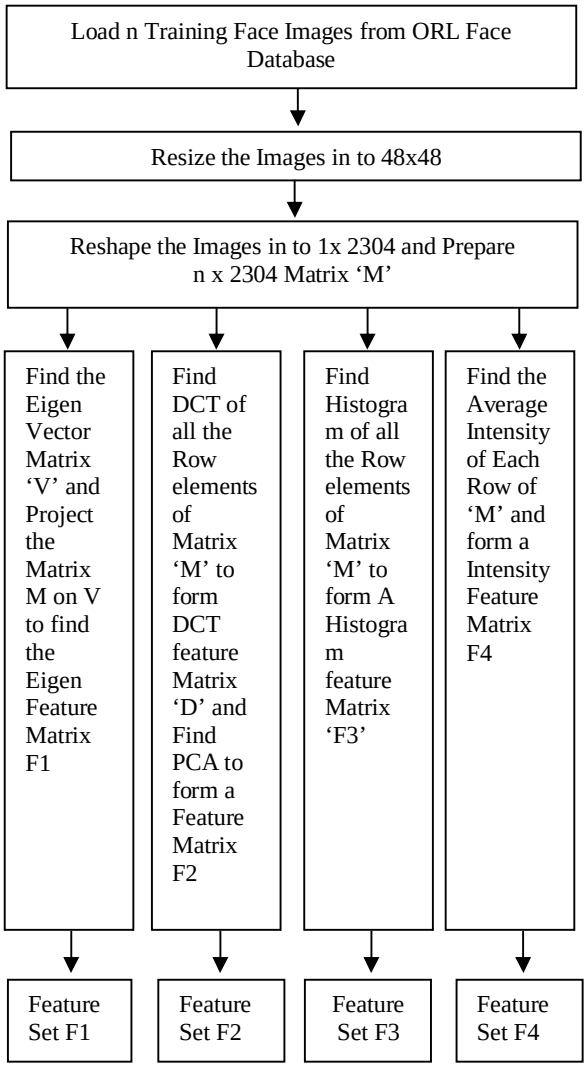


Figure 2: The Feature Set Creation

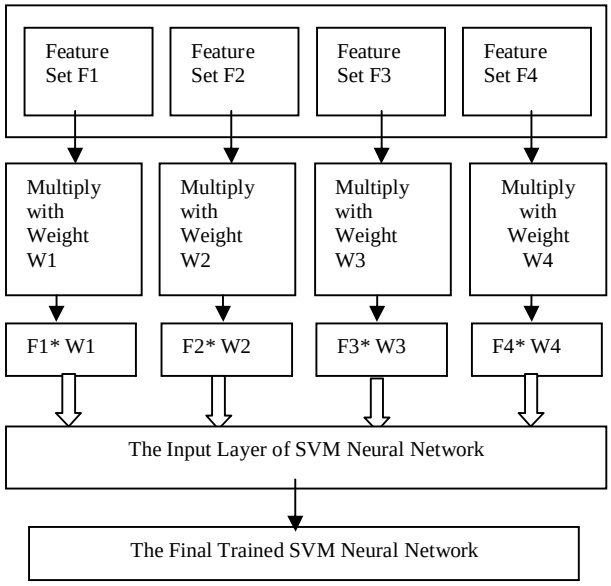


Figure 3: Architecture for Training and Testing the SVM with Weighted Attribute Set

A. *Steps Involved in Training*

1) Load a set of 'n' ORL Face Images For Training

2) Resize the Images in to 48x48 pixel size to reduce the memory requirements of the overall application

3) Reshape the Images in to 1x 2304 and prepare an n x 2304 Feature Matrix representing the training data set.

4) Apply a Feature Extraction, DCT, Histogram and Dimensionality Reduction technique and find the Four of above said Features sets F1, F2, F3 and F4.

5) Multiply the Weights W1, W2, W3 and W4 with F1, F2, F3 and F4.

6) Create an SVM network with "f1+f2+f3+f4" inputs where f1, f2, f3 and f4 are the corresponding feature lengths.

7) Train a SVM using the Weighted Feature Set.

B. *Steps Involved in Testing*

1) The first three steps of the above procedures will be repeated with test image set of ORL database to obtain a Feature Matrix representing the testing data set.

2) Project the matrix using Previous Eigen Vector Matrix and find the input Eigen Feature iF1.

3) Similarly find other input feature sets iF2, iF3 and iF4 of the input Image.

4) Classify the feature set [iF1 iF2 iF3 iF4] with previously trained SVM network.

5) Calculate Accuracy of Classification

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

IV. IMPLEMENTATION RESULTS AND ANALYSIS

The performance of proposed face recognition model was tested with the standard set of images called "ORL Face Database". The ORL Database of Faces contains a set of face images used in the context of a face recognition project carried out in collaboration with the Speech, Vision and Robotics Group of the Cambridge University Engineering Department. There are ten different images, each of 40 distinct subjects. For some subjects, the images were taken at different times, varying the lighting, facial expressions (open / closed eyes, smiling / not smiling) and facial details (glasses / no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position (with tolerance for some side movement).

Set of images from ORL databases were used for Training and Testing. The accuracy of recognition with multiple weighted attribute sets as well as a single attribute sets has been evaluated. The following table shows the overall results of these two types of techniques with different number of input face images.

TABLE 1 ACCURACY OF RECOGNITION WITH SINGLE ATTRIBUTE SET

No. of Faces used for Training and Testing	Accuracy of Recognition with different Weight Sets (%)			
	W1=1 W2=0 W3=0 W4=0	W1=0 W2=1 W3=0 W4=0	W1=0 W2=0 W3=1 W4=0	W1=0 W2=0 W3=0 W4=1
10	90.00	40.00	80.00	90.00
20	80.00	25.00	60.00	65.00
30	83.33	16.67	56.67	76.67
40	80.00	12.50	52.50	70.00
Average	**83.33**	**23.54**	**62.29**	**75.42**

TABLE 2 ACCURACY OF RECOGNITION WITH MULTIPLE ATTRIBUTE SET

No. of Faces used for Training and Testing	Accuracy of Recognition with different Weight Sets (%)				
	W1=1 W2=1 W3=0 W4=0	W1=.5 W2=1 W3=0 W4=0	W1=.5 W2=1 W3=0 W4=1	W1=1 W2=1 W3=1 W4=1	W1=.12 W2=0 W3=1 W4=0
10	90.00	100	100	100	100
20	80.00	85.00	85.00	85.00	90.00
30	86.67	86.67	86.67	83.33	86.67
40	82.50	85.00	82.50	82.50	82.50
Average	**84.79**	**89.17**	**88.54**	**87.71**	**89.79**

As shown in the above table and the following graphs, the performance of recognition while using Single Attribute Set as well as multiple attribute sets with same priority or weight will lead to poor recognition results. For example, all the results of **Table 1** which used a single attribute at a time for recognition, is in some what poor than **Table 2** which is using combined multi attributes. Further, if we note the fourth row (40 images) corresponding to the weights (W1=1, W2=1, W3=1, W4=1) which is using all the attributes with same weight, the result is poor while comparing it with column 2 (W1=0.5, W2=1, W3=0, W4=0). The following Graphs show the performance of the two different approaches. In **Figure 4** Line charts shows the Performance with single attribute set, but in **Figure 5** Line charts shows the performance with multiple weighted attribute sets.

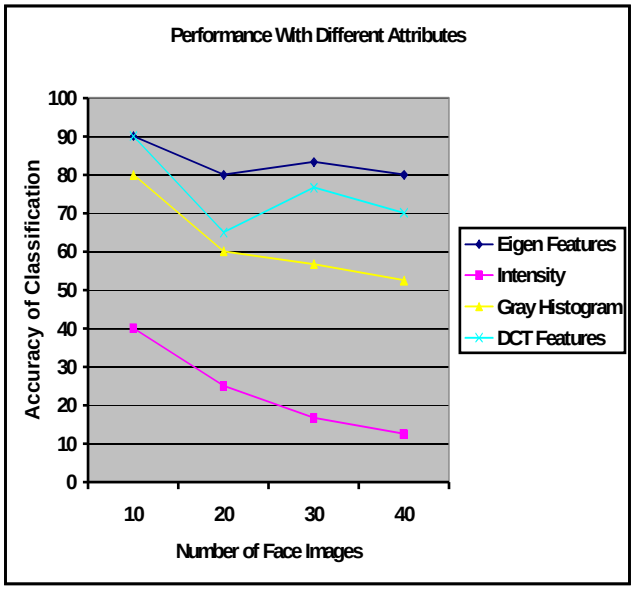


Figure 4 Performance with single attribute

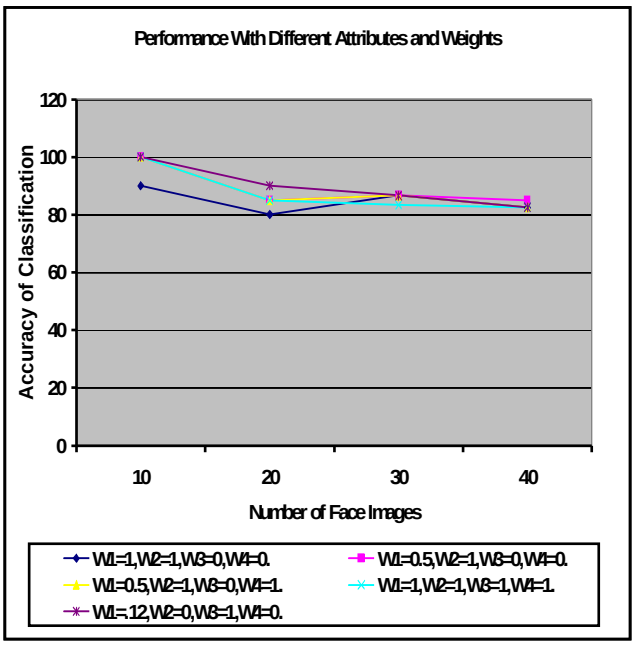


Figure 5 Performance with multiple weighted attribute sets

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

The Following Two Charts shows the average performance of the two approaches. In Figure 6, column charts shows the average performance with single attribute set.

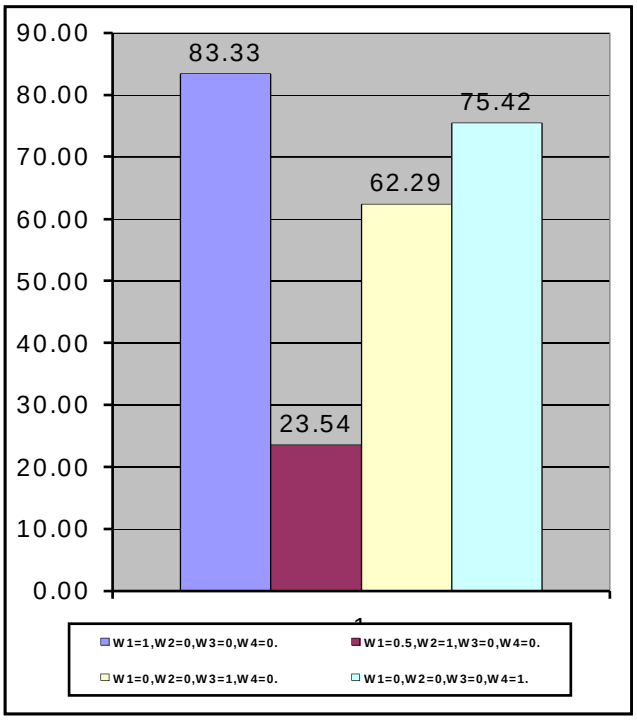


Figure 6 average performance with single attribute set

In **Figure 7,** column charts shows the average performance with multiple attribute sets.

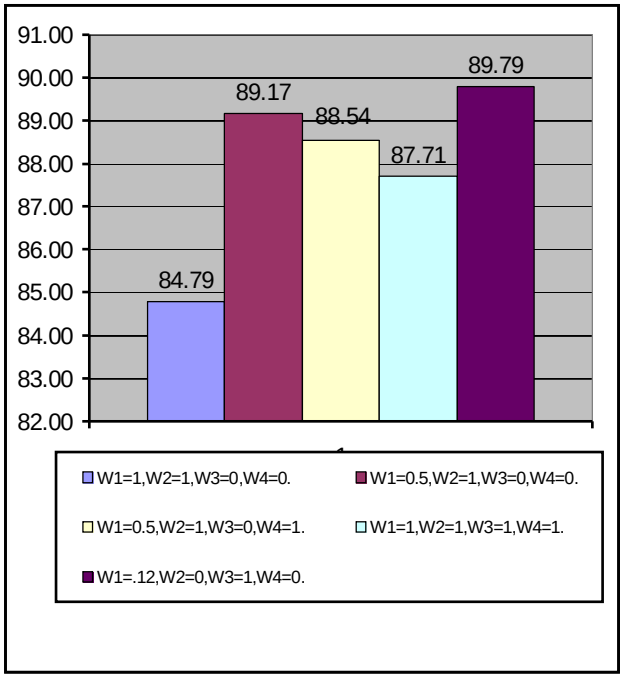


Figure 7 Average performances with multiple attribute set

V. CONCLUSION

Complicated Face recognition techniques constantly facing very challenging and difficult problems in several research works. In spite of the great work done in the last 30 years, it is sure that the face recognition research community will have to work for at least the next few decades to completely solve the problem. Strong and coordinated efforts between the computer visions signal processing and psychophysics and neurosciences community is needed. The proposed Multiple Weighted Feature Attribute Sets based training provided significant improvement in terms of performance accuracy of the face recognition system. With ORL data set, a significant 5% performance improvement was observed during various tests. In this work, we have selected PCA as the main feature extraction technique. The weights of the used feature sets were decided based on trial and error method. This will be applicable for systems with predefined data sets. In future works, one may explore different techniques like Kernel PCA, LDA for better performance. So, future works may address methods for automatic estimation of the weights of the feature sets with respect to the application.

REFERENCES

[1] S.Sakthivel, Dr.R.Lakshmipathi and M.A.Manikandan “Evaluation of Feature Extraction and Dimensionality Reduction Algorithms for Face Recognition using ORL Database”, The Paper published in Proceedings of The 2009 International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV'09) , Las Vegas, USA, 2009. ISBN: 1-60132-117-1, 1-60132-118-X (1-60132-119-8) Copyright 2009 CSREA Press.

[2] S.Sakthivel, Dr.R.Lakshmipathi and M.A.Manikandan “Improving the performance of machine learning based face recognition algorithm with Multiple Weighted Facial Attribute Sets” The paper published in proceedings of The Second International Conference on the Applications of Digital Information and Web Technologies, 2009 Volume, Issue, 4-6 Aug. 2009 Page(s):658 – 663, ISBN: 978-1-4244-4456-4, INSPEC Accession Number: 10905880, Digital Object Identifier: 10.1109/ICADIWT.2009.5273884

[3] Heng Fui Liau, Kah Phooi Seng, Li-Minn Ang and Siew Wen Chin, “New Parallel Models for Face Recognition” University of Nottingham Malaysia Campus, Malaysia - Recent Advances in Face Recognition, Published by In-Teh, 2008 Recent Advances in Face Recognition, Book edited by: Kresimir Delac, Mislav Grgic and Marian Stewart Bartlett, ISBN 978-953-7619-34-3, pp. 236, December 2008, I-Tech, Vienna, Austria.

[4] Khalid Youssef and Peng-Yung Woo, "A Novel Approach to Using Color Information in Improving Face Recognition Systems Based on Multi-Layer Neural Networks", Northern Illinois University USA - Recent Advances in Face Recognition, Published by In-Teh, 2008, Source: Recent Advances in Face Recognition, Book edited by: Kresimir Delac, Mislav Grgic and Marian Stewart Bartlett, ISBN 978-953-7619-34-3, pp. 236, December 2008, I-Tech, Vienna, Austria

[5] Xiaofei He; Deng Cai; Shuicheng Yan; Hong-Jiang Zhang,"Neighborhood preserving embedding", Tenth IEEE International Conference on Computer Vision, Volume 2, Issue, 17-21 Oct. 2005. Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV’05) 1550-5499/05

[6] Shuicheng Yan; Dong Xu; Benyu Zhang; Hong-Jiang Zhang,"Graph embedding: a general framework for dimensionality reduction", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Volume 2, June 2005. Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05) 1063-6919/05.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

[7] S. Canu, Y. Grandvalet, V. Guigue, and A. Rakotomamonjy. (2005). SVM and kernel methods Matlab toolbox Perception Systems at Information http://asi.insa-uen.fr/~arakotom/toolbox/index.html

[8] Xiaofei He, Partha Niyogi, "Locality Preserving Projections (LPP)", Computer Science Department, The University of Chicago, In Advances in Neural Information Processing Systems, Volume 16, page 37, Cambridge, MA, USA 2004, The MIT press. http://books.nips.cc/papers/files/nips16/NIPS2003_AA20.pdf

[9] Lu, J.; Plataniotis, K.N. & Venetsanopoulos, A. N. (2003) Face Recognition Using LDA based Algorithm", *IEEE trans.Neural Network*, vol.14, No 1, pp.195-199, January 2003. Digital Object Identifier 10.1109/ TNN.2002.806647.

[10] Yu, Hu. & Yang, J. (2001) A Direct LDA algorithm for high-dimension data with application to face recognition, *Pattern Recognition*, vol.34, pp. 2067-2070, 2001. DOI: 10.1016/S0031-3203(00)00162-X

[11] Chen, L.F.; Mark Liao, H.Y.; Ko, M.T.; Lin, J.C. & Yu, G.J. (2000) A new LDA-based face recognition system which can solve the small space size problem, *Pattern Recognition*, vol.33, pp.1703-1726, 2000. DOI: 10.1016/ S0031-3203(99)00139-9

[12] H.A. Rowley and T. Kanade, "Neural network-based face detection", IEEE Trans. on PAMI, vol. 20, no. 1, pp. 23-38,Jan1998.http://www.informedia.cs.cmu.edu/documents/ rowley-ieee.pdf

[13] E. Osuna, R. Freund, and F. Girosi, "Training support vector machines: an application to face detection," in Proc. IEEE Conf. Computer Vision Pattern Recognition, 1997. http://doi.ieeecomputersociety.org/10.1109/CVPR.1997.609310

[14] Imola K. Fodor, Center for Applied Scientific Computing, Lawrence Livermore National Laboratory "A survey of dimension reduction techniques"June2002. https://computation.llnl.gov/casc/sapphire/pubs/148494.pdf

[15] Turk, M. & Pentland, A. (1991) Eigenfaces for recognition, *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, Mar 1991. http://portal.acm.org/ citation.cfm?id=1326894

[16] Kirby, M. & Sirovich, L. (1990) Application of the Karhunen-Loeve procedure of the characteristic of human faces, *IEEE Trans. Pattern Anal. Machine Intell*, vol.12, pp 103-108, Jan, 1990. DOI 10.1109/34.41390

[17] Shang-Hung Lin, "An Introduction to Face Recognition Technology", Information science special issue on multimedia informing Technologies - part2 volume 3 No 1, 2000. http://inform.nu/Articles/Vol3/v3n1p01-07.pdf

AUTHORS PROFILE

[1] **S. Sakthivel** received his M.E Computer science from Sona College of Technology, Affiliated to Anna University, Chennai, India in the year 2004.Currently, pursuing his Ph.D., in Anna University, Chennai,Tamilnadu.He has a work experience of 10 years. At present working as a Assistant Professorin the department of Information Technology. He has published paper in international journal.He has participated and presented research papers in various national and international seminars and conferences. He is an Life member of ISTE.

[2] **Dr.R.Lakshimpathi** received the B.E degree in 1971 and M.E degree in 1973 from College of Engineering, Guindy, and Chennai.He received his PhD degree in High Voltage Engineering from Indian Institute of Technology, Chennai, India. He has 36 years of teaching experience in various Government Engineering Colleges in Tamilnadu and he retired as Principal and Regional Research Director at Alagappa Chettiar College of Engineering and Technology, Karaikudi.He is now working as Professor in Electrical and Electronics Engineering department at St.Peters Engineering College, Chennai.His areas of research include HVDC Transmission, Power System Stability and Electrical Power Semiconductor Drives.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

A DNA and Amino Acids-Based Implementation of Playfair Cipher

Mona Sabry(1), Mohamed Hashem(2), Taymoor Nazmy(1),
Mohamed Essam Khalifa(3)
Faculty of Computer Science and information systems,
Ain Shams University,
Cairo, Egypt.
E-mail: mona.sabry@hotmail.com

ABSTRACT-- The DNA cryptography is a new and very promising direction in cryptography research. Although in its primitive stage, DNA cryptography is shown to be very effective. Currently, several DNA computing algorithms are proposed for quite some cryptography, cryptanalysis and steganography problems, and they are very powerful in these areas.

This paper discusses a significant modification to the old Playfair cipher by introducing DNA-based and amino acids-based structure to the core of the ciphering process.

In this study, a binary form of data, such as plaintext messages, or images are transformed into sequences of DNA nucleotides. Subsequently, these nucleotides pass through a Playfair encryption process based on amino-acids structure.

The fundamental idea behind this encryption technique is to enforce other conventional cryptographic algorithms which proved to be broken, and also to open the door for applying the DNA and Amino Acids concepts to more conventional cryptographic algorithms to enhance their security features.

KEY WORDS: DNA, amino acids, encryption, decryption, cryptography, security, Playfair cipher.

I. INTRODUCTION

As some of the modern cryptography algorithms (such as DES, and more recently, MD5) are broken, the new directions of information security are being sought to protect the data. The concept of using DNA computing in the fields of cryptography and steganography is a possible technology that may bring forward a new hope for powerful, or even unbreakable, algorithms.

The main purpose behind our work is to discover new fields of encoding the data in addition to the conventional used encryption algorithm in order to increase the concept of confusion and therefore increase security.

In our work, we applied the conversion of character form or binary form of data to the DNA form and then to amino acid form. Then the resulting form goes through the encryption algorithm which we chose for example; the classical Playfair cipher.

It is Adleman, with his pioneering work [5]; set the stage for the new field of bio-computing research. His main idea was to use actual chemistry to solve problems that are either unsolvable by conventional computers, or require an enormous amount of computation. By the use of DNA computing, the Data Encryption Standard (DES) cryptographic protocol can be broken [6]. In DNA steganography, A DNA encoded message is first camouflaged within the enormous complexity of human genomic DNA and then further concealed by confining this sample to a microdot[3]. Recent research considers the use of the Human genome in cryptography. In 2000, the Junior Nobel Prize was awarded to a young Romanian American student, Viviana Risca, for her work in DNA steganography.[3]

The one-time pad cryptography with DNA strands, and the research on DNA steganography (hiding messages in DNA), are shown in [2] and [3].

However, researchers in DNA cryptography are still looking at much more theory than practicality. The constraints of its high tech lab requirements and computational limitations, combined with the labor intensive extrapolation means. Thus prevent DNA computing from being of efficient use in today's security world.

Another approach is lead by Ning Kang in which he did not use real DNA computing, but just used the principle ideas in central dogma of molecular biology to develop his cryptography method. The method only simulates the transcription, splicing, and translation process of the central dogma; thus, it is a pseudo DNA cryptography method. [4]

There is another investigation conducted by [1] which is based on a conventional symmetric encryption algorithm called "Yet Another Encryption Algorithm" (YAEA) developed by Saeb and Baith [1]. In this study, he introduces the concept of using DNA computing in the fields of cryptography in order to enhance the security of cryptographic algorithms. This is considered a pioneering idea that stood behind our work in this paper. **[**1**]**

(1) Computer Science Dept. (2) Information Systems Dept. (3) Basic Science Dept.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Although Playfair cipher is believed to be an old, simple and an easily breakable cipher, we believe our new modifications can make it a more powerful encryption algorithm. This is done by introducing concepts of confusion and diffusion to the core of the encryption process in addition to preserving the cipher's simplicity concept.

In addition shortage in security features the plaintext message is restricted to be all upper case, without J letters, without punctuation, or even numerical values. Those problems can be easily handled in any modern cipher as handled in our new algorithm [8].

The character form of a message or any form of an image can be easily transformed to the form of bits. This binary form can be transformed to DNA form through many encoding techniques implemented in previous work and summarized in [7].

Playfair is based on the English alphabetical letters, so preserving this concept, we will use the English alphabet but from an indirect way. DNA contains four bases that can be given an abbreviation of only four letters (adenine (A), cytosine (C), guanine (G) and thymine (T)). On the other side, we have 20 amino acids with additional 3 codons to represent the Stop of coding region. Each amino acid is abbreviated by a single English character. So we are able to stretch these 20 characters to 26 characters, we will be able to represent the English alphabet.

Then, we have to convert the DNA form of data to amino acid form so that it can go through a classical Playfair cipher. Through this conversion process, we have to keep in mind the problem of ambiguity; that most amino acids are given more than possible codon.

The rest of the paper is organized as follows: section two will give a brief explanation of what is DNA and process of Transcription and translation. Section three introduces our new algorithm followed by a detailed explanation of the encryption/decryption processes. Section four shows the experiment steps, results. Section five introduces some additional security features. Section six shows conclusion and future work.

II. OVERVIEW OF DNA

A. What is Deoxyribonucleic acid 'DNA'?

DNA is a nucleic acid that contains the genetic instructions used in the development and functioning of all known living organisms and some viruses. The main role of DNA molecules is the long-term storage of information. DNA is often compared to a set of blueprints or a recipe, or a code, since it contains the instructions needed to construct other components of cells, such as proteins and RNA molecules. The DNA segments that carry this genetic information are called genes, but other DNA sequences have structural purposes, or are involved in regulating the use of this genetic information.

The DNA double helix is stabilized by hydrogen bonds between the bases attached to the two strands. The four bases found in DNA are adenine (abbreviated A), cytosine (C), guanine (G) and thymine (T). These four bases are attached to the sugar/phosphate to form the complete nucleotide, as shown for adenosine monophosphate.

B. The genetic code

The genetic code consists of 64 triplets of nucleotides. These triplets are called **codons**. With three exceptions, each codon encodes for one of the 20 amino acids used in the synthesis of proteins. That produces some redundancy in the code: most of the amino acids being encoded by more than one codon.

The genetic code can be expressed as either RNA codons or DNA codons. RNA codons occur in messenger RNA (**mRNA**) and are the codons that are actually "read" during the synthesis of polypeptides (the process called **translation**). But each mRNA molecule acquires its sequence of nucleotides by **transcription** from the corresponding gene.

The DNA Codons is read the same as the RNA codons Except that the nucleotide thymidine (**T**) is found in place of uridine (**U**). So in DNA codons we have (TCAG) and in RNA codons, we have (UCTG).

C. Transcription and translation

A gene is a sequence of DNA that contains genetic information and can influence the phenotype of an organism. Within a gene, the sequence of bases along a DNA strand defines a messenger RNA sequence, which then defines one or more protein sequences. The relationship between the nucleotide sequences of genes and the amino-acid sequences of proteins is determined by the rules of translation, known collectively as the genetic code. The genetic code consists of three-letter 'words' called codons formed from a sequence of three nucleotides (e.g. ACT, CAG, TTT).

In transcription, the codons of a gene are copied into messenger RNA by RNA polymerase. This RNA copy is then decoded by a ribosome that reads the RNA sequence by base-pairing the messenger RNA to transfer RNA, which carries amino acids. Since there are 4 bases in 3-letter combinations, there are 64 possible codons (4^3 combinations). These encode the twenty standard amino acids, giving most amino acids more than one possible codon. There are also three 'stop' or 'nonsense' codons signifying the end of the coding region; these

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

are the TAA, TGA and TAG codons.RNA codon table, WIKIPEDIA: http://en.wikipedia.org/wiki/Genetic_code#cite_note-pmid19056476-8

III. DNA-BASED PLAYFAIR ALGORITHM

A. *Encryption algorithm of DNA-based Playfair cipher:*

Playfair used to be applied to English alphabet characters of plaintext. It was unable to encode any special characters or numbers which is considered a severe drawback that enforces the sender to write everything in the English letters. This problem appears while sending numerical data, equations or symbols.

On the contrary, in our algorithm, we can use any numbers, special characters or even spaces (not preferred) in or plaintext. The encryption process starts by the binary form of data (message or image) which is transferred to DNA form according to Table 1. Then the DNA form is transferred to the Amino acids form according to Table 2 which is a standard universal table of Amino acids and their codons representation in the form of DNA [RNA codon table, Wikipedia: http://en.wikipedia.org/wiki/Genetic_code#cite_note-pmid19056476-8].

Note that each amino acid has a name, abbreviation, and a single character symbol. This character symbol is what we will use in our algorithm.

Table I: DNA Representation of bits.

Bit 1	Bit 2	DNA
0	0	A
0	1	C
1	0	G
1	1	T

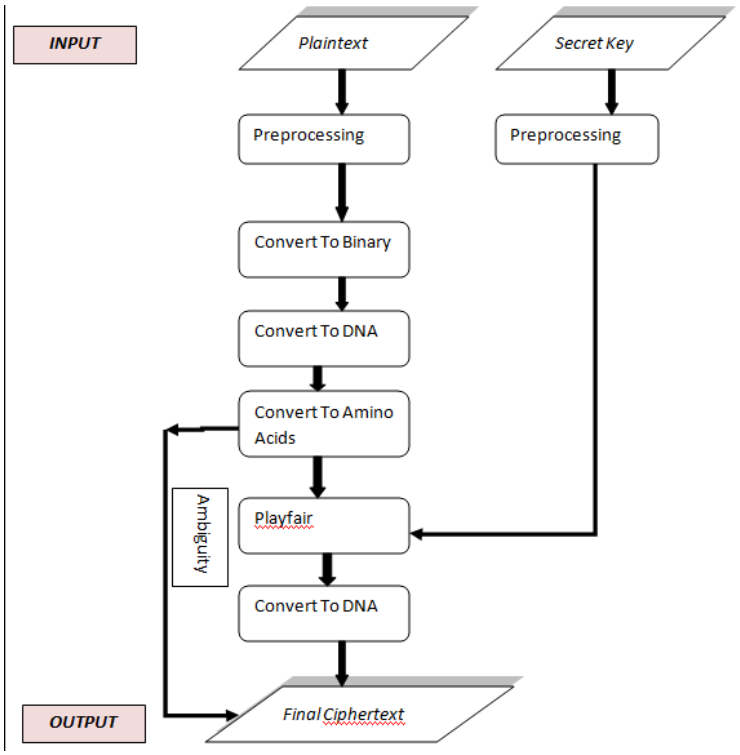


Fig. 1: flowchart of the DNA-based Playfair algorithm

B. *Constructing the alphabet table:*

In the table, we have only 20 amino acids in addition to 1 start and 1 stop. While we need 25 letters to construct the Playfair matrix (note that I/J are assigned to one cell).

The letters we need to fill are (B, O, U, X, Z). So we will make these characters share some amino acids their codons.
The start codon is repeated with amino acid (M) so we will not use it. We will assign to (B) the 3 stop codons. We have 3 amino acids (L, R, S) having 6 codons. By noticing the sequence of DNA of each, we can figure out that each has 4 codons of the same type and 2 of another type. Those 2 of the other type are shifted to the letters (O, U, X) respectively. Letter (Z) will take one codon from (Y), so that Y: UAU, Z: UAC. Now the new distribution of codons is illustrated in Table 3.

Counting the number of codons of each character, we will find the number varies between 1 and 4 codons per character. We will call this number 'Ambiguity' of the character [AMBIG].

Now we have the distribution of the complete English alphabet, so a message in the form of Amino Acids can go through traditional Playfair cipher process using the secret key.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Table II: Amino acids and their 64 codons

Ala/A	GCU, GCC, GCA, GCG	**Leu/L**	UUA, UUG, CUU, CUC, CUA, CUG
Arg/R	CGU,CGC, CGA,CGG, AGA,AGG	**Lys/K**	AAA, AAG
Asn/N	AAU, AAC	**Met/M**	AUG
Asp/D	GAU, GAC	**Phe/F**	UUU, UUC
Cys/C	UGU, UGC	**Pro/P**	CCU,CCC, CCA, CCG
Gln/Q	CAA, CAG	**Ser/S**	UCU, UCC, UCA, UCG, AGU, AGC
Glu/E	GAA, GAG	**Thr/T**	ACU, ACC, ACA, ACG
Gly/G	GGU,GGC, GGA,GGG	**Trp/W**	UGG
His/H	CAU, CAC	**Tyr/Y**	UAU, UAC
Ile/I	AUU, AUC, AUA	**Val/V**	GUU, GUC, GUA, GUG
START	AUG	**STOP**	UAA, UGA, UAG

Table III: New distribution of the alphabet with the corresponding new codons:

	STOP										from			To			from	from		to			to	from	To
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
4	3	2	2	2	2	4	2	3		2	6	1	2		4	2	6	6	4		4	1		2	
GCU	UAA	UGU	GAU	GAA	UUU	GGU	CAU	AUU		AAA	UUA	AUG	AAU	UUA	CCU	CAA	CGU	UCU	ACU	AGA	GUU	UGG	AGU	UAU	UAC
GCC	UAG	UGC	GAC	GAG	UUC	GGC	CAC	AUC		AAG	UUG		AAC	UUG	CCC	CAG	CGC	UCC	ACC	AGG	GUC		AGC	UAC	
GCA	UGA					GGA		AUA			CUU				CCA		CGA	UCA	ACA		GUA				
GCG						GGG					CUC				CCG		CGG	UCG	ACG		GUG				
											CUA						AGA	AGU							
											CUG						AGG	AGC							
4	3	2	2	2	2	4	2	3	3	2	4	1	2	2	4	2	4	4	4	2	4	1	2	1	1

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Table IV: New Distribution for codons on English alphabet .I

A	GCU, GCC, GCA, GCG	**L**	CUU, CUC, CUA, CUG
R	CGU, CGC, CGA, CGG	**K**	AAA, AAG
N	AAU, AAC	**M**	AUG
D	GAU, GAC	**F**	UUU, UUC
C	UGU, UGC	**P**	CCU,CCC, CCA, CCG
Q	CAA, CAG	**S**	UCU, UCC, UCA, UCG
E	GAA, GAG	**T**	ACU, ACC, ACA, ACG
G	GGU,GGC, GGA,GGG	**W**	UGG
H	CAU, CAC	**Y**	UAC
I	AUU, AUC, AUA	**V**	GUU, GUC, GUA, GUG
B	**UAA, UGA, UAG**	**O**	**UUA, UUG**
U	**AGA,AGG**	**X**	**AGU, AGC**
Z	**UAC**		

The output form is the amino acid form of cipher text. DNA form of cipher text can be demonstrated also from Table 4 by choosing random codons accompanied to each character. The concept that one character can have more than one DNA representation is itself an addition to confusion concept that enhances the algorithm strength. Table\ IV shows the new distribution of codons on the amino acids and additional alphabetical English letters according to our algorithm.

C. Decryption and Ambiguity problem

The decryption process is simply the inverse of the encryption process unless that we will find a problem in constructing the DNA form of plaintext from the amino acid form which is of length (L). The problem is that we are unable to choose which codon to put in accordance to each amino acid character. This is simply the problem of codon-amino acid mapping problem arised with other algorithm based on the concept of Central Dogma like [4]. The way Nang handled this problem is to put this codon-amino acid mapping in the secret key to be sent through a secure channel [4]. This idea is not efficient since it increases the size of the key in relation to size of the plaintext.

The solution in our algorithm is located in two additional bits for each amino acid character to demonstrate which codon to choose.
We said before that each amino acid has 1, 2, 3 or 4 codons to represent it. This is a number that can be put in 2 bits from 0→3.
These 2 bits can be converted to DNA form from Table 1.

That is why the final cipher text is both the DNA form of cipher text of length (3L) and the array carrying the ambiguity of length (L).
In decryption, the amino acid form of plaintext with the assistance of the ambiguity array can construct the correct form of plaintext in DNA form which can be transferred to binary form and then the final character form.

D. Pseudo-code

Input:

[P] Plaintext (characters with spaces, numbers or any special characters).

[K] Secret key (English characters without any number or special characters).

Algorithm body:

Preprocessing:

1- Prepare the secret key:
- Remove any spaces or repeated characters from [K].
- Put the remaining characters in the UPPER case form. [K]→UPPER[K].

2- Prepare the plaintext:
- Remove the spaces from [P] (done to avoid attacker's trace to a character which is repeated many times within the message)

Processing:

1- Binary form [BP] = BINARY [P] (Replace each character by its binary representation-8 bits-)
2- DNA form [DP] = DNA [BP] (Replace each 2 bits by their DNA representation)
3- Amino acids form [AP] = AMINO [DP] (Replace each 3 DNA characters by their Amino acid character keeping in track the ambiguity of each Amino acid [AMBIG].
4- Construct the Playfair 5X5 matrix and add [K] row by row, then add the rest of alphabet characters.
5- Amino acid of cipher text [AC]= PLAYFAIR [AP].
6- DNA form of cipher text [DC] = DNA [AC].

Output:

Add [DC] and [AMBIG] together in the suitable form→ final cipher text [C].

E. Samples of the program steps and output

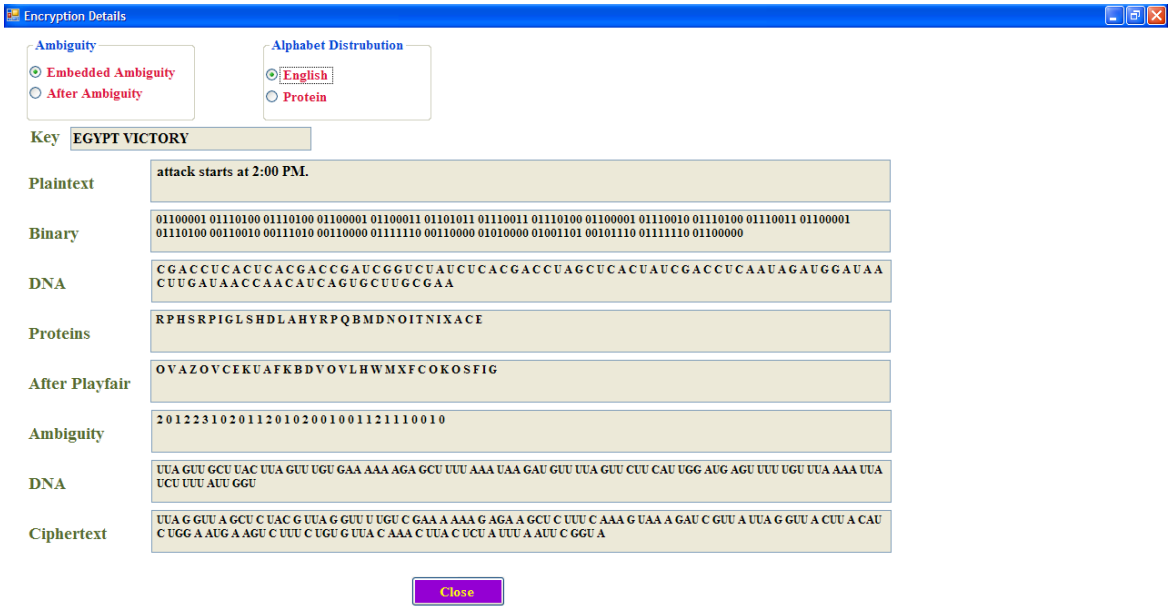


Figure 2: sample of steps of encryption implementation

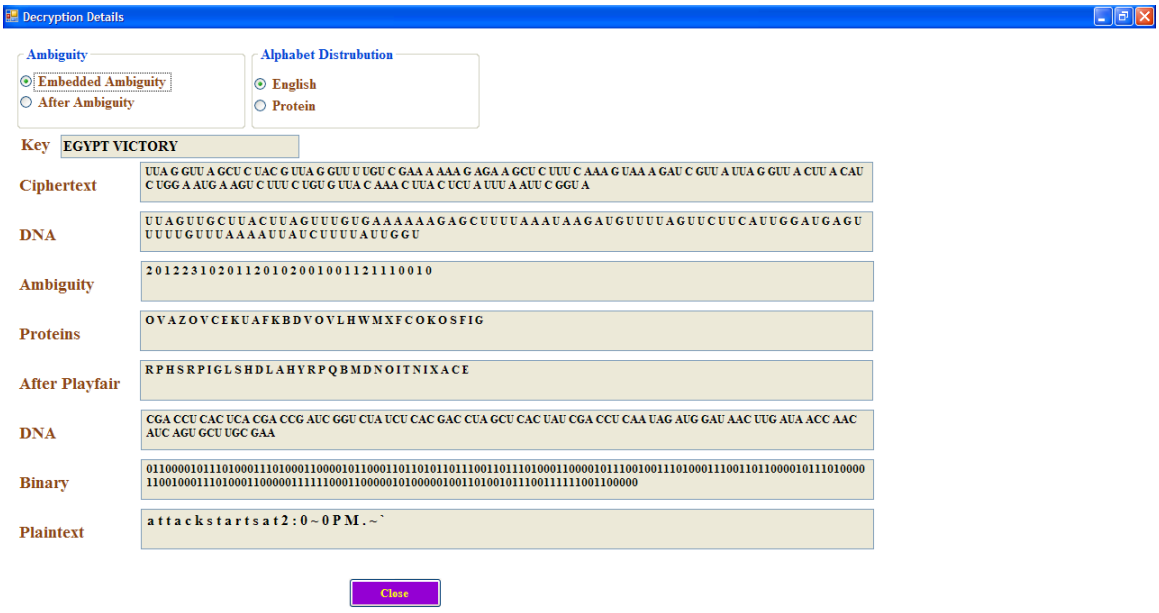


Figure 3: sample of steps of decryption implementation

IV. EXPERIMENT AND PERFORMANCE ANALYSIS

A. *Experiment*

1- Experiment inputs and attributes

We led our experiment on the famous novel 'A Tale of Two Cities' by Charles Dickens found on http://www.literature.org/authors/dickens-charles/two-cities/ . We will take paragraphs from the beginning of the novel according to the estimated storage size in Kilobytes (from 1 KB and increasing till 150 KB).

2- System Parameters

The experiments are conducted using Intel(R) Core (TM) 2CPU T5300, 1.73 GHz, 32 bit processor with 1GB of RAM. The simulation program is compiled using the default settings in .NET 2005 visual studio for C# windows applications under WINDOWS XP as the operating system. The experiments will be performed several times to assure that the results are consistent and valid.

3- Experiment Factors

The chosen factor here to determine the performance is the algorithm's speed to encrypt data blocks of various sizes. Suppose we will use the original sequence of English alphabet and embed the ambiguity inside the message not after it. The secret key used is "CHARLES DICKENS" which results in 11Bytes key.

4- Experiment steps:

Experiment preprocessing:

1- Loading the table of the 64 amino acids with their DNA Encodings and number of ambiguous encodings.
2- Formatting the secret key by removing spaces, repeated characters and non English letters.
3- Formatting the plaintext by removing spaces between words and separating the repeated doubles by the character '~' which chosen to be a rarely used character.

Processing:

This includes:
1- Converting characters to binary form.
2- Converting binary to DNA
3- Converting the DNA to amino acids and recording ambiguity.
4- Do Playfair encryption.
5- Convert the amino acid form of cipher text to DNA form in addition to embedding the ambiguity in the DNA format.

5- Experiment Results

The next table illustrates the experiments and time taken to encrypt each piece of plaintext (each is of different data loads) in milliseconds.
The time taken by loading the amino acids table and preparing the secret key is ignored because it is comparatively small to processing time.

Table 4: performance results of DNA-based Playfair algorithm

Input size of plaintext (in KB)	Plaintext after preprocessing	Preprocessing plaintext	From Binary to Amino Acids form	playfair	Prepare ciphertext	Total processing time	Bytes/Second
1 (1,022 B)	846B	0	0	0	15.625	15.625	65.408
10 (9,757 B)	8124B	62.500	15.625	0	125.000	203.125	48.034
20 (20,023 B)	16599B	203.125	15.625	0	171.875	390.625	51.259
50 (50,432 B)	41781B	1062.500	46.875	15.625	437.500	1562.500	32.276
100 (97,072 B)	83910B	4687.500	78.125	31.25	859.375	5656.250	17.162
150 (153,418 B)	127098 B	11390.625	140.625	31.25	1343.750	12906.25	11.887

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

V. Additional security features

We have illustrated the main core of the algorithm and now we are going to suggest some additional features to the algorithm which can enhance its security and strength.

A- ***The key*****:**
It is quite clear that the more random and long the key is, the more the difficulty to break the cipher will be.

B- ***Use Amino acids alphabet sequence instead of English alphabetical sequence*****:**
The standard table of amino acids has a special sequence defined in the matrix [4X4] (UCAG) X (UCAG). This sequence of acids can be used instead of the sequence of English alphabet letters to fill the rest of the 5X5 matrix after adding the secret key.

C- ***Combine the total resulting message into long strand of DNA to be inserted in a microdot (steganography):***
One of the advantages of this algorithm is the variety of ways we can use to write down the cipher text. It can be written in DNA form, binary form or even character form which is more confusing. The advantage of DNA form is that it can make use of several steganography techniques developed for DNA messages [3]. It can also be prepared in biological labs like in [2] in which DNA message goes through a biological DNA encryption process using one time pad or substitution.

D- ***Ambiguity a problem that contains useful confusion feature*****:**
Some characters in table 2 can have 6 codons representing the problem of ambiguity. The way we handled the preparation of table 3 made each character have in maximum 4 codons. The number 4 can be represented by 2 bits and therefore ban be represented by one DNA character. That was a benefit that made us able to write the cipher text with ambiguity in the form of DNA.

E- ***Use of conventional XOR-ing procedure*****:**
Another way to increase security is defining another key that can be XOR-ed with the amino acid form or DNA form of cipher text. It was a pioneer idea by [1] to choose the key as the DNA strand of a certain organism. This idea assures the key randomness and variety in length according to the length of the message.

VI. CONCLUSION AND FUTURE WORK

The fundamental idea behind this technique is to open the door for the idea of applying the DNA and Amino Acids encoding concepts to other conventional cryptographic algorithms to enhance their security -vulnerability- features.
Our algorithm initially succeeded in overcoming some main problems in "Playfair cipher" like restriction of plaintext to "English Alphabet". As in our algorithm the plaintext is to be converted to its binary value before encryption, it now clear that the plaintext message can be written in upper or lower case, with any punctuation, and numerical values.

Other papers conducted the idea of amino acids way of representation from the point of view of the central dogma design [4]. But they were unable to clearly handle the problem of ambiguity as performed by our algorithm. Our algorithm made few preprocessing steps to handle this problem and the result was quite accurate (same input message obtained after decryption). This feature is very important when regarding an encryption algorithm in order to verify the concept of data integrity or in other words, to assure that data after decryption to be the same input data before encryption.
Finally, our algorithm provides different forms of the cipher text like: Binary form, DNA form, Amino Acid form or character form. Those various forms can match different used applications.

Our future work is dedicated to implementing this encoding on other known algorithms and measuring its performance and security, Also, Experiments should be conducted to implement the algorithm on different applications to ensure its feasibility and applicability.

REFERENCES

[1] Sherif T. Amin, Magdy Saeb, Salah El-Gindi, **"A DNA-based Implementation of YAEA Encryption Algorithm,"** IASTED International Conference on Computational Intelligence (CI 2006), San Francisco, Nov. 20, 2006. http://www.actapress.com/PaperInfo.aspx?PaperID=29058.

[2] Ashish Gehani, Thomas LaBean and John Reif. DNA-Based Cryptography. DIMACS DNA Based Computers V, American Mathematical Society, 2000.

[3] TAYLOR Clelland Catherine, Viviana Risca, Carter Bancroft, 1999, "**Hiding Messages in DNA Microdots".** Nature Magazine Vol.. 399, June 10, 1999.

[4] KANG Ning, **"A Pseudo DNA Cryptography Method**", Independent Research Study Project for CS5231, October 2004.

[5] Leonard Adleman. "**Molecular Computation of Solutions to Combinatorial Problems**". Science, 266:1021-1024, November 1994.

[6] Dan Boneh, Cristopher Dunworth, and Richard Lipton. **"Breaking DES Using a Molecular Computer**". Technical Report CS-TR-489-95, Department of Computer Science, Princeton University, USA, 1995.

[7] Dominik Heider and Angelika Barnekow**, "DNA-based watermarks using the DNA-Crypt algorithm"**, Published: 29 May 2007 BMC Bioinformatics 2007, **8**:176 doi:10.1186/1471-2105-8-176, http://www.biomedcentral.com/1471-2105/8/176 ,© 2007 Heider and Barnekow; licensee BioMed Central Ltd.

[8] William Stallings. "**Cryptography and Network Security"**, Third Edition, Prentice Hall International, 2003.

Ultra Wideband Slot Antenna with Reconfigurable Notch bands

J. William* and R.Nakkeeran
Department of Electronics and Communication Engineering
Pondicherry Engineering College
Puducherry, India . 605014.
Email id: wills.susan@gmail.com, rnakeeran@pec.edu

***Corresponding author:** wills.susan@gmail.com

***Abstract*— An Ultra Wideband (UWB) slot antenna with reconfigurable notch bands is presented in this paper. The basic UWB antenna consists of a rectangular slot with triangular structure that acts as a tuning stub with CPW feed. The CPW feed is designed for 50 Ω impedance. The notch band is achieved by inserting a rectangular slot in the ground plane with a effective length of λ/2. The reconfigurable rejection of the bands 3.1 GHz – 3.9 GHz, 4 GHz–5.3 GHz and 4.1 GHz–5.9 GHz are achieved by switching the diodes placed over the slot in the ground plane. The characteristics of the designed structure are investigated by using MoM based electromagnetic solver, IE3D. The return loss (S_{11}) of the antenna is measured and that are comparable with the simulation results. The proposed antenna covers the entire UWB range 3.1 GHz to 10.6 GHz with reconfiguration. The low profile and simple configuration of the proposed antenna leads to easy fabrication that may be built in any wireless UWB device applications where reconfigurable rejection bands are required. The rejection of WiMax, IEEE 802.11a and HYPERLAN/2 bands can be achieved by using the proposed antenna design.**

Keywords- ***coplanar waveguide; notch band; slot antenna; reconfiguration; ultra wideband***

I. INTRODUCTION

In the year 2002, Federal Communications Commission (FCC) released the unlicensed UWB spectrum 3.1 GHz to 10.6 GHz for the commercial purposes. After the release of UWB, it gains much attention by the researchers due to its inherent properties of low power consumption, high data rate and simple configuration [1]. With the rapid developments of such UWB systems, a lot of attention is being given for designing the UWB antennas. Designing an antenna to operate in the UWB band is quiet challenging one because it has to satisfy the requirements such as ultra wide impedance bandwidth, omni directional radiation pattern, constant gain, high radiation efficiency, constant group delay, low profile, easy manufacturing etc [2]. Interestingly the planar slot antennas with CPW fed posses the above said features with simple structure, less radiation loss, less dispersion and easy integration of monolithic microwave integrated circuits (MMIC) [3]. Due to inherent nature of UWB system sharing the same spectrum with other systems. The interference from other system should be considered when you design a UWB system. For example, the other wireless communication systems such as WiMax (3.3 GHz – 3.7 GHz), IEEE 802.11a (5.15 GHz - 5.35 GHz and 5.725 GHz – 5.825 GHz) and HYPERLAN/2 (5.15 GHz - 5.35 GHz and 5.47 GHz – 5.825 GHz) which are also operated in the portion of UWB band. Therefore, the UWB antenna design could play an important role when interference with other wireless applications such as WiMax, IEEE 802.11a and HYPERLAN/2 systems which are also coexisting with UWB band (3.1 GHz – 10.6 GHz), which degrades the UWB system performance. Therefore, a UWB antenna having reconfigurable frequency band notch characteristics is enviable.

To mitigate the above electromagnetic interference from nearby communication systems many antenna designs have been proposed in the literature [4-13]. Many techniques also used to introduce notch band for rejecting the interference in the UWB slot antennas. It is done either by inserting half wavelength slits, stripes in the tuning stub [14], or inserting stub in the aperture connected to the ground planes [15], or inserting square ring resonator in the tuning stub [16], or inserting 'L' branches in the ground plane [17], or with complementary split ring resonator [18], or inserting strip in the slot [19]. All the above methods are used for rejecting a fixed band of frequencies. But to effectively utilize the UWB spectrum and to improve the performance of the UWB system, it is desirable to design the UWB antenna with reconfigurable notch band. It will help to minimize the interference between the systems and to improve the performance of the UWB systems. In general, reconfiguration are popular in antenna engineering for their frequency agility, bandwidth enhancement and polarization diversity [20-23]. In [24] RF MEMS switches are used for the reconfigurabilty of rejection band between the frequencies 5 GHz to 6 GHz. In this paper a new method is proposed to obtain the reconfiguration in the frequency notch bands for WiMax, WLAN 802.11a and HYPERLAN/2 systems by making 'on' and 'off' the diodes placed over the slot that is introduced on the ground plane. This intern changes the tuning length of the slot, which is responsible for the desired frequency notch band. Through our simulation it is also found that the desired frequency notch band can be obtained by inserting slot in the tuning stub with appropriate length rather than on the ground plane. However this paper mainly focuses

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

on reconfigurable notch band through the introduction of slot in the ground plane. The simulation software used for this analysis is IE3D [25]. The paper is organized as follows: Section II brings out the geometry of the antenna. In section III, simulation results and analysis are presented. Obtained experimental results are given in Section IV. Section IV concludes the work.

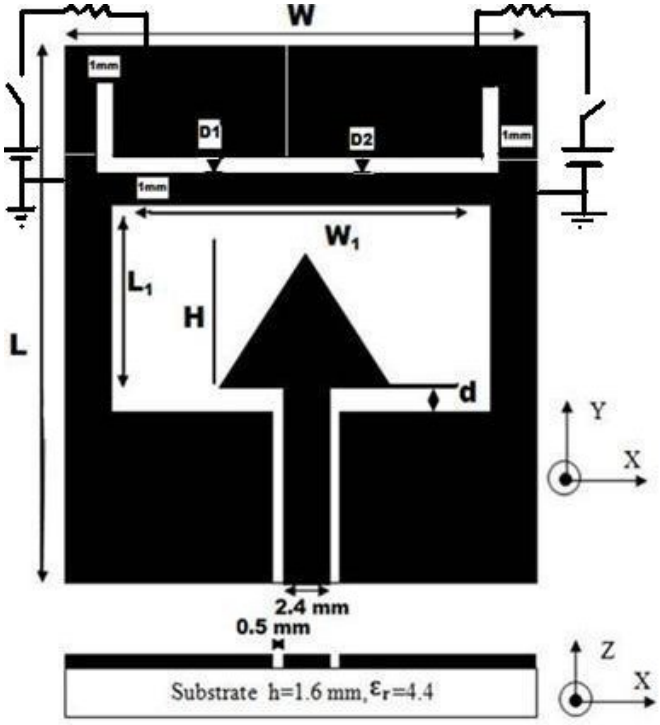


Fig. 1 Geometry of the proposed slot antenna with biasing circuit.

II. Antenna Geometry

The structure of the antenna is shown in Fig 1. The antenna consists of rectangular slot with width 'W_1' and length 'L_1'. The tuning stub comprises a triangular patch with height 'H'. The distance between the tuning stub and feed line is 'd', 'W' and 'L' are the overall width and length of the antenna respectively. In this study, the dielectric substance (FR4) with thickness of 1.6 mm with relative permittivity of 4.4 is chosen as substrate to facilitate printed circuit board integration. The CPW feed is designed for 50 Ω characteristic impedance with fixed 2.4 mm feed line width and 0.5 mm ground gap. Slits are introduced to avoid the short circuit between the ground planes. Fast switching diode LL4148 is used as a switching device. The proposed antenna is designed to cover the entire UWB band with reconfiguration capability. The placement of the diodes 'L_d' are desired by the effective wavelength λ_{eff},

The effective wavelength of the slot is,

$$\lambda_{\text{eff}} = \frac{c}{f_n\sqrt{\varepsilon_{eff}}}, \quad \varepsilon_{eff} \approx \frac{\varepsilon_r + 1}{2} \tag{1}$$

The placement of the diodes 'L_d'

$$L_d = \frac{c}{2f_n\sqrt{\varepsilon_{eff}}} \tag{2}$$

where 'f_n' is the centre frequency of the notch band. The placement of the diodes is desired by the effective wavelength for the different notch frequencies.

III. Simulated Results and Analysis

The analysis and performance of the proposed antenna is explored by using IE3D for the better impedance matching. The detailed parametric analysis of the UWB antenna is carried out and presented in our paper [26]. In order to evaluate the performance of the proposed antenna with reconfigurable slot in the ground plane, the optimal parameter values of the antenna (without slot) suggested in that paper are considered. The final optimal parameter values of the antenna are listed in the Table 1. However to study the impact of the slot, the slot length 'L_2' and width 'W_2' are varied by keeping one of them as constant. In the simulation switching diodes were simulated as capacitor for the 'Off' state and as a resistor in the 'On' state. The current distribution, gain and group delay are also studied. The simulated return loss, with different states of the diode are shown in Fig. 2, clearly indicate that the notch frequency and notch bandwidth is varied when the slot length is varied.

TABLE I. Optimal Parameter Values of the Antenna

Parameter	Description	Optimal Value
L	Length of the antenna	28 mm
W	Width of the antenna	21 mm
L_1	Length of the slot	15 mm
W_1	Width of the slot	16.8 mm
d	Feed gap distance	1.6 mm
H	Height of the patch	9.3 mm

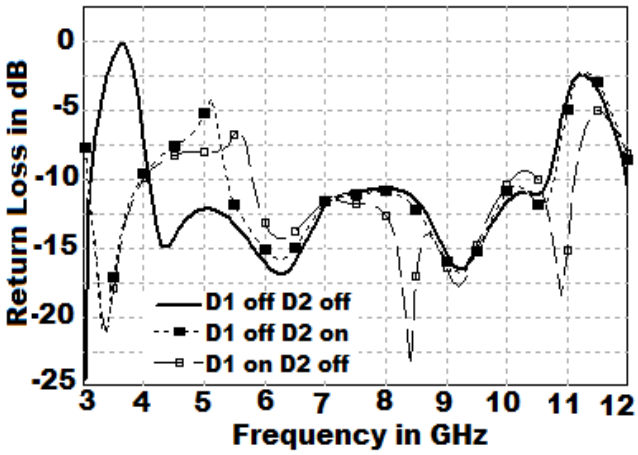


Fig. 2 Simulated response curves for different biasing condition.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

A. Effect of Slot in the Ground Plane

The length of the ground slot is adjusted by switching the diodes placed appropriately over the slot. The switching diodes 'D_1' and 'D_2' are placed over the slot along the ground plane. The Table 2 shows the states of the diode switches and notch frequency with notch band width.

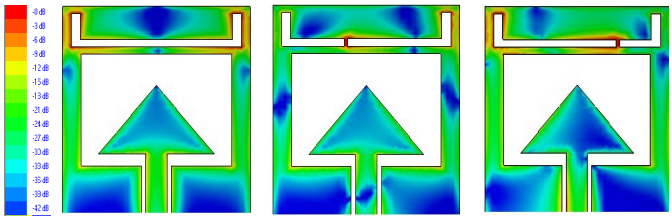
TABLE II. DESCRIPTION OF SWITCHING OF THE DIODES

D_1	D_2	Notch Freq. (GHz)	Notch band
Off	Off	3.6 GHz	3.1-3.9 GHz
Off	On	5.1 GHz	4 – 5.3 GHz
On	Off	5.6 GHz	4.1-5.9 GHz

When both diodes are in open state the effective slot length is longer to achieve the notch frequency at 3.6 GHz with frequency notch band of 3.1 GHz to 3.9 GHz, which is the WIMAX band of frequency. When D_1 is 'Off' and D_2 is 'On' the length of the slot is tuned to eliminate the band of 4 GHz to 5.3 GHz with notch center frequency of 5.1 GHz which covers the lower band of IEEE 802.11a. When D_1 is 'On' and D_2 is 'Off' the effective wavelength of the slot is adjusted such that to eliminate the upper band of IEEE 802.11a with a notch frequency of 5.6 GHz with the notch frequency band of 4.1 GHz to 5.9 GHz.

B. Simulated Current Distribuition

The simulated current distribution at different notch frequencies according to the switching condition of the diodes is presented in Fig.3, it is witnessed from the figure, the current distribution around the radiation slot is disturbed by the introduction of ground slot, which is responsible for the notch in the frequency band. If the slot length is longer the destructive interference with the main radiating slot is high which causes the notching in that band is good. When the ground slot length is decreased, the current distribution of the radiating slot is disturbed by the tuning slot is less effective compared with the longer slot size.



(a) (b) (c)

Fig. 3 Simulated current distribution for different switching conditions of diodes. a) 'D_1 and 'D_2' 'Off' b) 'D_1' 'On' and 'D_2' 'Off"' c) 'D_1' 'Off' and 'D_2' 'On'

C. Gain

The computed gain of the proposed UWB antenna for different tuning lengths of the slot in the ground plane and without slot is compared in Figure 4.

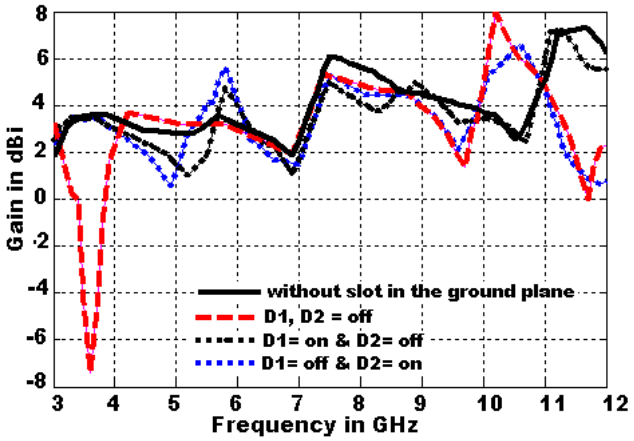
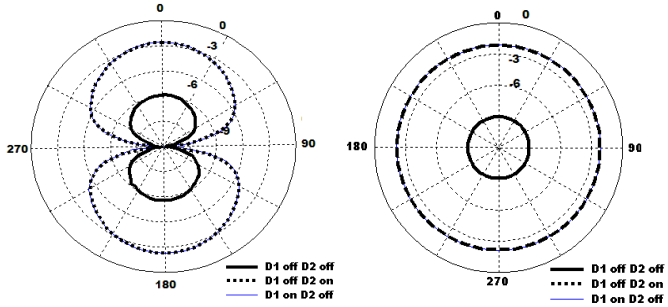


Fig. 4 Simulated gain for different biasing conditions of diodes.

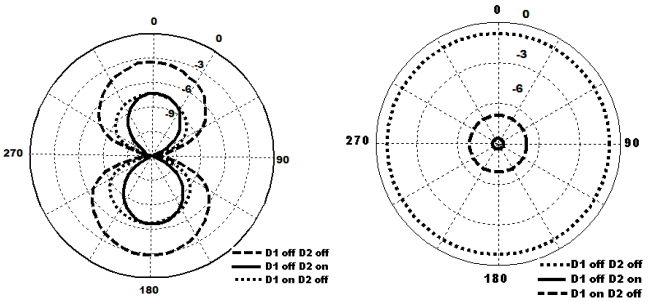
It is observed that there is a gain variation at the notched frequencies when the length of the slot is tuned. which is -7.5 dBi for both diodes are 'off' state and 0.4 dBi when D_1 is 'Off' and D_2 is 'On'. Where as it is 0.6 dBi for D_1 is 'On' and D_2 is 'Off'. The antenna without slot and the gain varies from 2 dBi to 6 dBi across the UWB spectrum.

D. Radiation pattern

The radiation pattern for the E plane and H plane at frequencies 3.4 GHz, 5.2 GHz and 5.6 GHz are simulated for all the three cases of diode states are compared and displayed in Fig.5, which disclose that the directivity gain of the radiation pattern is reduced at the notch frequencies without affecting the shape of the radiation pattern. In the E plane, it is bidirectional pattern and in H plane, it is omni directional pattern.

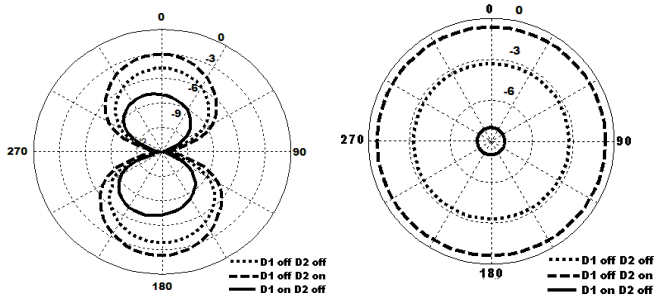


(a) 3.4 GHz



(b) 5.2 GHz

http://sites.google.com/site/ijcsis/
ISSN 1947-5500



(c) 5.6 GHz

Fig. 5 Simulated E-plane and H-plane radiation patterns at notch frequencies.

E. Group delay

The group delay 'τ' of the antenna is calculated from the phase of the computed 'S_{21}' by using the following equation and plotted in Fig. 6,

$$\tau = - \frac{d\phi}{df} \qquad (2)$$

where 'ϕ' is phase of S_{21} in radians /sec and 'f' is frequency in GHz. From the Fig. 6, it is noticed that the variation in the group delay for the antenna with and without slot is around 2 ns for the frequency range from 3.1 GHz to 10.6 GHz. There is a variation in the group delay at the notch band in the response which is due to notch behavior of the antenna.

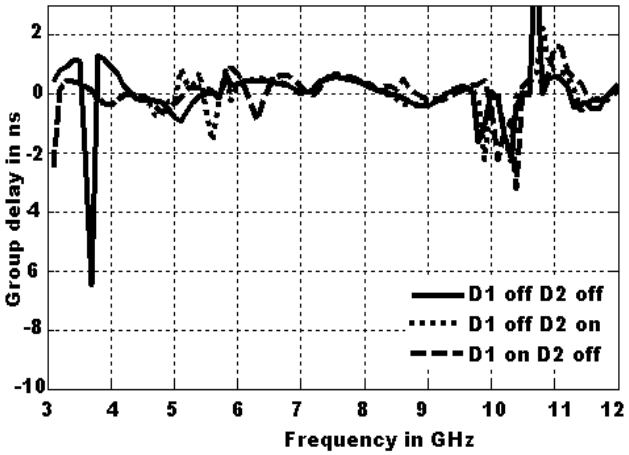


Fig. 6 Group delay response.

IV. Experimentall Results and Discussion

The prototype of the proposed antenna shown in Fig. 1 was fabricated for different parameters with their optimal values and tested. Using Hewlett Packard Network Analyzer (HP8757D), the VSWR is measured and compared with the simulation result is shown in Fig.7. There is a discrepancy between the measured and the simulated ones might be the effect of soldering the SMA connector or fabrication tolerance. The simulation result was obtained by assuming coplanar as input port, whereas practically SMA connector was used, the imperfect transition between SMA feed to coplanar may introduce losses [27] and also the capacitances can lead to shift in the frequency.

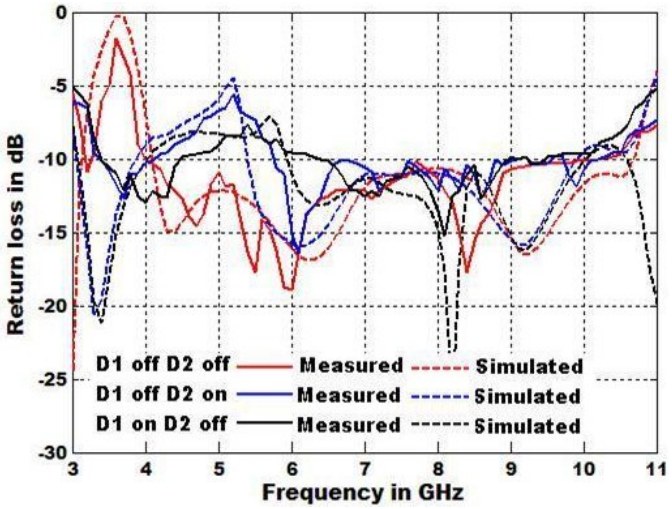
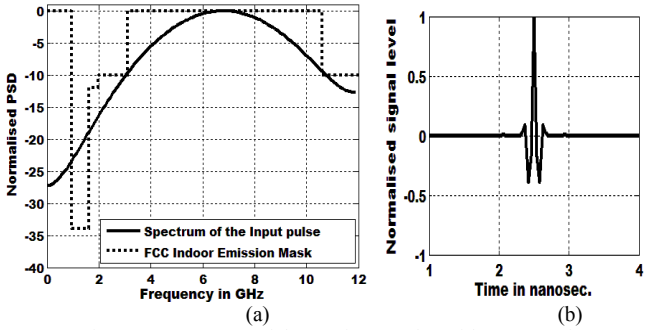


Fig. 8 Comparison of measured and simulated VSWR.

V. Time Domain Analysis

In ultra wideband systems, the information is transmitted using short pulses. Hence, it is important to study the temporal behavior of the transmitted pulse. The communication system for UWB pulse transmission must provide as minimum as possible distortion, spreading and disturbance. The channel is assumed to be linear time invariant (LTI) system to verify the capability of the proposed antenna for transmission and reception of these narrow pulses. The transfer function of the entire system is computed using simulated value of 'S_{21}' parameter [28]. The received output pulse is obtained by taking the Inverse Fourier Transform (IFT) of the product of transfer function and spectrum of the test input pulse. While computing 'S_{21}', two identical antennas are placed face to face at a distance of 75 cm that is greater than the far-field distance of the antenna. The cosine modulated Gaussian pulse is considered for this analysis with centre frequency of 6.85 GHz and pulse width of 220 picoseconds, whose spectrum is shown in Fig.7. It satisfies the requirement of FCC mask for UWB indoor emission. The comparison of input and output responses of the system for the antenna with different notches are shown in Fig.8, which ensures the distortion less pulse transmission and also guarantees that the designed antenna is capable of transmitting and receiving short pulses.



(a) (b)

Fig. 7. a) Spectrum of the test input pulse with FCC mask
b) input pulse in time domain

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

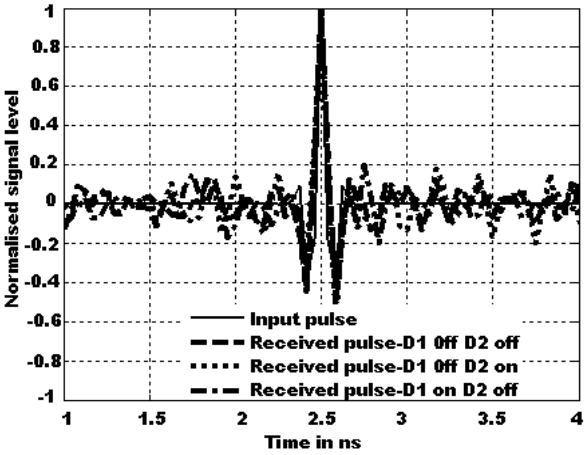


Fig. 8. Comparison of input pulse and received pulses in time domain.

VI. CONCLUSION

This paper describes the simulated analysis of UWB slot antenna with reconfigurable rejection bands. By switching the diodes, the effective length of the ground slot is tuned to the notch center frequency. The simple configuration and low profile of the proposed antenna leads to easy fabrication that may be built for any wireless UWB device applications where reconfigurable rejection of notch bands is required. The rejection of WiMax, IEEE 802.11a and HYPERLAN/2 bands can be achieved using the proposed antenna.

REFERENCES

[1] FCC NEWS(FCC 02-48), Feb. 14,2002. FCC News release.

[2] M. Ghavami, L.B. Michael and R. Kohno, *Ultra Wideband Signals and Systems in Communication Engineering*, New York: John Wiley and Sons, USA, 2004.

[3] K.L. Wong, *Compact and Broadband Microstrip Antenna*, John Wiley and sons. Inc., NY, USA, 2001.

[4] I.-J. Yoon, H. Kim et al, "Ultra-wideband tapered slot antenna with band cutoff characteristic," *Electronics Letters*, vol. 41, no. 11, pp. 629-630, May 2005.

[5] A. M. Abbosh, M. E. Bialkowski et al, " A Planar UWB antenna With Signal Rejection Capability in the 4–6 GHz Band," *IEEE Microwave and Wireless Comp. Letters*, vol. 16, no. 5, pp. 278-280, May 2006.

[6] Keyvan Bahadori and Yahya Rahmat-Samii, "Miniaturized Elliptic-Card UWB antenna with WLAN Band Rejection for Wireless Communications," *IEEE Trans. Antennas and Propag*., vol.55, no.11, pp.3326-3332,Nov. 2007.

[7] Chong-Yu Hong, Ching-Wei Ling et al, " Design of a Planar Ultra-wideband Antenna With a New Band-Notch Structure ," *IEEE Trans.Antennas and Propag*. vol.55, no.12, pp.3391-3397, Dec. 2007.

[8] Q.-X. Chu and Y.-Y. Yang , " 3.5 / 5.5 GHz dual band-notch ultra-wide band antenna," *Electronics Letters,* vol. 44, no. 3, pp. 172-174, Jan. 2008.

[9] L. Wang, . Wu, X.-W. Shi et al , "Design Of A Novel monopole UWB Antenna with a notched ground," *PIER*, vol. 5, pp. 13–20, 2008.

[10] Jia-Yi Sze and Jen-Yi Shiu, " Design of Band-Notched Ultrawideband square aperture antenna with a hat-Shaped Back-Patch," *IEEE Trans. Antennas and Propag*., vol.56, no.10, pp.3391-3397,Oct. 2008.

[11] S. Zhou, J. Ma and J. Deng, " A Novel dual band-notched ultra wideband," Journal of Electromagn. Waves and Appln., vol. 23, pp. 57-63, 2009.

[12] J.Y. Deng, Y.Z. Yin, J. Ma and Q.Z.Liu, "Compact Ultra –wideband with dual band-notched characteristics," Journal of Electromagnetic Waves and Appln., vol. 23, pp. 109-116, 2009.

[13] X. Li, L.Yang, S.X. Gong and Y.J Yung, " A Novel tri band-notched monopole antenna," Journal of Electromagnetic Waves and Appln., vol. 23, pp. 139-147, 2009.

[14] Yi-Cheng Lin and Kuan-Jung Hung, "Compact Ultrawideband Rectangular Aperture Antenna and Band-Notched Designs," *IEEE Trans. Antennas Propag.,* vol.54, no.11, pp.3075-3081, Nov.2006.

[15] Hyung Kuk Yoon, Yohan Lim et al, " UWB Wide Slot antenna with Band -notch Function," *Proc. IEEE Antennas and Propag. society Intl Symposium*, pp. 3059-3062, July 2006.

[16] Wen-jun Lui, Chong-hu Cheng and Hong-bo Zhu, "Improved Frequency Notched Ultra wideband Slot Antenna Using Square Ring Resonator," *IEEE Trans. Antennas Propag.,* vol.55, no.9, pp.2445-2450, Sept.2007.

[17] Yunlong Cai and Zhenghe Feng , " A UWB Antenna with novel L branches on ground for Band-Notching Application ," *Proc. Of IEEE Intl. Conference on Microwave and Millimeter wave Tec*., vol. 4, pp.1654-1657, April 2008.

[18] The-Nan Chang and Min-Chi Wu , " Band-Notched Design for UWB Antennas , " *IEEE Antennas Wirel. Propag. Letters*, vol.7, pp.636- 639, 2008.

[19] C.- Y. Huang, S.- A. Huang and C.- F. Yang , " Band-notched ultra-wideband circular slot antenna with Inverted C-shaped parasitic strip,"*Electronics Letters* , vol. 44, no.15, pp. 891-892, July 2008.

[20] Abdel-Fattah Sheta and Samir F. Mahmoud, "A Widely Tunable Compact Patch Antenna, " *IEEE Antennas Wirel. Propag. Letters*, vol.7, pp.40- 42, 2008.

[21] M.-I. Lai,T. Y . W u,J.-C. Hsieh,C.-H. Wang and S.-K. Jeng, "Design of reconfigurable antennas based on an L-shaped slot and PIN diodes for compact wireless devices," IET Microw. Antennas Propag., vol. 3, no. 1, pp. 47 – 54, 2009.

[22] Julien Sarrazin , Yann Mahé, Stéphane Avrillon, and Serge Toutain, "Pattern Reconfigurable Cubic Antenna," *IEEE Trans. Antennas Propag.,* vol.57, no.2, pp.310-317, Feb. 2009.

[23] Yevhen Yashchyshyn, Jacek Marczewski, Krzysztof Derzakowski, Jozef W. Modelski , and Piotr B. Grabiec, "Development and Investigation of an Antenna System With Reconfigurable Aperture," *IEEE Trans. Antennas Propag.,* vol.57, no.1, pp.2-8, Jan. 2009.

[24] Symeon Nikolaou Nickolas D. Kingsley , George E. Ponchak, John Papapolymerou, and Manos M. Tentzeris, "UWB Elliptical Monopoles with a Reconfigurable Band Notch Using MEMS Switches Actuated without Bias Lines," *IEEE Trans. Antennas Propag.,* vol.57, no.8, pp.2242-2250, Aug. 2009.

[25] IE3D 14, Zeland Software, Ins., Fremont, USA.

[26] J. William and R. Nakkeeran , "A new compact CPW-fed wideband slot antenna for UWB applications," *Proc. of IEEE First Himalayan Intl. Conference on Internet* , Nov. 2009. OI 10.1109/ AHICI.2009.5340282.

[27] Kuang-ping ma, Yongxi Qian and Tatsuo Itoh, "Analysis and applications of a New CPW- slot line Transition," *IEEE Transactions on Microwave theory and Techniques*, vol. 47, pp. 426-432, April 1999.

[28] Stanislas Licul and William A Davis, "Ultra-wideband(UWB) antenna measurements using vector Network analyzer," *IEEE Antennas and Propagation International Symposium,* pp. 1319-1322, 2004.

AUTHORS PROFILE

J.William received his B.E. degree in Electronics and Communication from Bharathidasan University, Tamilnadu, India, and the M.Tech.degree in Communication Systems from National Institute of Technology (N.I.T), Trichy,India, in 1991 and 2006 respectively. He is currently working towards the Ph.D. degree at Pondicherry Engineering College, Pondicherry, India. He is a life member of ISTE and IE (I) and member of IEICE and EurApp. His current research interest is in the area of coplanar waveguide feed antennas and printed slot antennas for UWB.

R. Nakkeeran Received BSc. Degree in Science and B.E degree in Electronics and Communication Engineering from the Madras University in 1987 and 1991 respectively and M.E degree in Electronics and Communication Engineering (diversification in Optical Communication) from the Anna University in 1995. He received Ph.D degree from Pondicherry University in 2004. Since 1991, he has been working in the teaching profession. Presently, he is Assistant Professor in Pondicherry Engineering College. He is life member of IETE, ISTE, OSI and IE(I). Also he is member

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

of OSA, SPIE and IEEE. He has published seventy five papers in National and International Conference Proceedings and Journals. He has co-authored a book, published by PHI. His areas of interest are Optical Communication, Networks, Antennas, Electromagnetic Fields and Wireless Communication.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

UWB Slot Antenna with Rejection of IEEE 802.11a Band

J. William* and R.Nakkeeran
Department of Electronics and Communication Engineering
Pondicherry Engineering College
Puducherry, India . 605014.
Email id: wills.susan@gmail.com, rnakeeran@pec.edu

***Corresponding author:** wills.susan@gmail.com

Abstract— **A compact coplanar waveguide (CPW) fed slot antenna for ultra wideband (UWB) with notched band from 5.1 GHz to 5.9 GHz is presented in this paper. By inserting a rectangular slot with the particular length and width in the ground plane, a desired notch in the frequency band can be achieved. The characteristics of the designed structure are investigated using an electromagnetic solver, IE3D. The overall size of the antenna comes around 28(L) × 21(W) × 1.6(T) mm^3. For the developed antenna the VSWR is measured and compared with the simulated results. The measured parameter is in good agreement with the simulation and the antenna covers entire UWB band ranging from 3.1 GHz to 11.4 GHz with band notching between 5.1 GHz and 5.9 GHz. Time domain analysis of the antenna is also investigated and presented, which ensures that the antenna is capable of working effectively in the UWB environment. This type of antenna configuration would be quiet useful for UWB indoor applications with no interference from WLAN and HYPERLAN/2 systems when they coexist.**

Keywords- ***band notch; coplanar waveguide; slot antenna; time domain analysis; ultra wideband***

I. INTRODUCTION

To support variety of applications to the mobile users needs wireless communication systems with higher capacity and data rate. Ultra wideband (UWB) is a short pulse communication system offers both of the above requirements. After the release of UWB by Federal Communications Commission (FCC), it has become one of the interesting technologies in indoor wireless communications system and receives much attention by the industries and academia due to its properties of low power consumption, support of high secured data rate and simple configuration [1]. Designing an antenna to operate in the UWB band is quiet challenging one because it has to satisfy the requirements such as ultra wide impedance bandwidth, omni directional radiation pattern, constant gain, high radiation efficiency, constant group delay, low profile, compact and easy manufacturing [2]. Interestingly the planar slot antennas with CPW fed possesses the above said features with simple structure, less radiation loss, less dispersion and easy integration of monolithic microwave integrated circuits (MMIC) [3]. Most of the UWB antennas have a wide bandwidth, covering bands used for other wireless communication applications. Therefore, the UWB antenna design could play an important role when interference with other wireless applications such as IEEE 802.11a and HYPERLAN/2 systems which are also coexisting with UWB band (5.1 GHz - 5.9 GHz), which degrades the UWB system performance.

Therefore, a UWB antenna having frequency band notch characteristics is enviable. To mitigate this electromagnetic interference from nearby communication systems many antenna designs have been proposed in the literature [4-10]. Many techniques also used to introduce notch band for rejecting the interference in the UWB slot antennas. It is done either by inserting half wavelength slits, stripes in the tuning stub [11], or inserting stub in the aperture connected to the ground planes [12], or inserting square ring resonator in the tuning stub [13], or inserting 'L' branches in the ground plane [14], or with complementary split ring resonator [15], or inserting strip in the slot [16]. In this paper, a new type of UWB antenna and its notched design is proposed by inserting a slot in the ground plane at a half wavelength size at the desired center notch frequency of 5. 5 GHz, which is an average of notch band 5.1 GHz to 5.9 GHz. The proposed antenna is simulated and analyzed by using simulation software, IE3D 14 [17]. The details of the proposed design and its experimental results are presented and discussed in the following sections. The paper is organized as follows: Section II brings out the design and geometry of the antenna. In Section III simulation results and analysis are presented. Obtained experimental results are given in Section IV. Section V explains time domain analysis of the antenna. Section VI concludes the paper.

II. ANTENNA GEOMETRY

In designing this type of antennas, the width 'W' and length 'L' play a crucial role in determining the resonant frequency of the system. The initial values of these parameters are calculated by using the equations given in [18] for the substrate height (h), dielectric constant (ε_r) and for the lower frequency. The designed values of the antenna are optimized with IE3D tool. The optimization was performed for the best impedance bandwidth. The proposed antenna with slot in the ground plane in this paper is a dimensional modification of structure given in [19]. The notched design for the particular

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

band is obtained by introducing a slot in the ground plane nearer to the radiating slot.

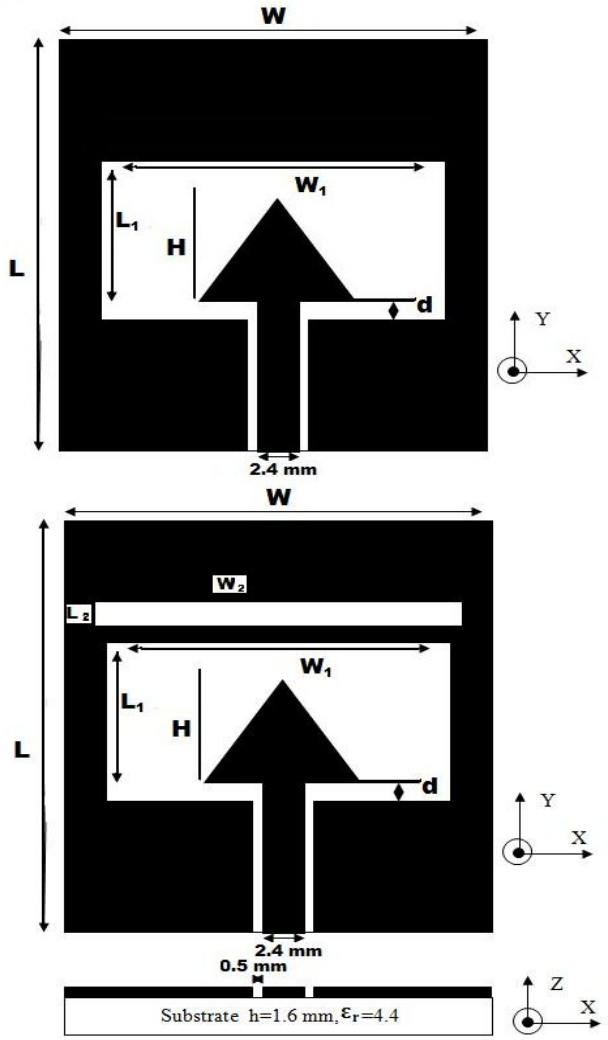


Fig. 1 Geometry of the proposed slot antenna and its notched design.

The effective length of the slot width is found to be approximately 0.5 λ_{eff} for the notch in the frequency band at the center frequency of 5.5 GHz. The introduction of the slot in the ground plane causes destructive interference in the current distribution which is responsible for the elimination of the specific frequency band.
The effective wavelength of the slot width is,

$$\lambda_{eff} = \frac{c}{f_n \sqrt{\varepsilon_{eff}}}, \quad \varepsilon_{eff} \approx \frac{\varepsilon_r + 1}{2} \qquad (1)$$

where 'f_n' is the centre frequency of the notch band.

The structure of the basic UWB slot antenna and its notched design is shown in Fig 1. The antenna consists of rectangular slot with width 'W_1' and length 'L_1'. The tuning stub comprises a triangular patch with height 'H'. The distance between the tuning stub and feed line is 'd'. 'W_2' and 'L_2' are the width and length of slot in the ground plane. 'W' and 'L' are the overall width and length of the antenna respectively. In this study, the dielectric substance (FR4) with thickness of 1.6mm with relative permittivity of 4.4 is chosen as substrate to facilitate printed circuit board integration. The CPW feed is designed for 50 Ω characteristic impedance with fixed 2.4 mm feed line width and 0.5 mm ground gap.

III. Simulated Results and Analysis

The analysis and performance of the proposed antenna is explored by using IE3D for the better impedance matching. The parametric analysis of the antenna carried out by varying one parameter and keeping other parameters constant. The simulated return loss of the proposed antenna is shown in Fig. 2, which clearly indicates that the impedance bandwidth of the antenna is 8.3 GHz (3.1GHz -11.4 GHz) for return loss (S_{11}) < -10 dB (VSWR < 2). The ultra wideband is due to multiple resonances introduced by the combination of the rectangular slot and the tuning stub. The resonant frequency and bandwidth are controlled by the size of the rectangular slot, antenna and tuning stub. Detailed parametric analysis has been carried out and the final optimal parameter values of the antenna are listed in the Table 1. However to study the impact of the slot, the slot length 'L_2' and width 'W_2' are varied by keeping one of them as constant. The current distribution, gain and group delay are also studied.

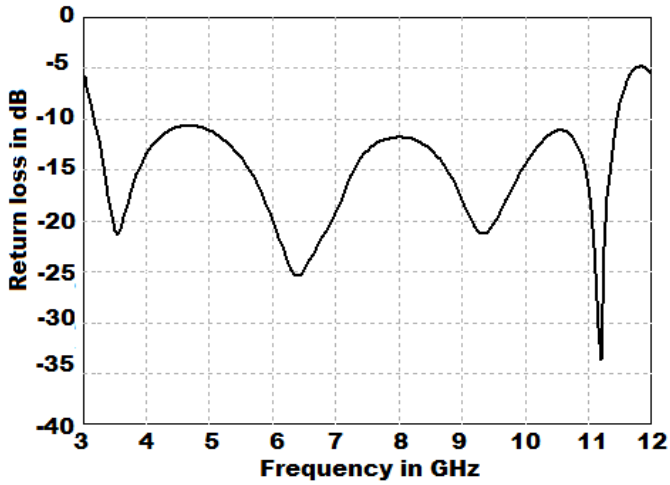


Fig. 2 Simulated return loss of the proposed UWB antenna.

TABLE I. Optimal Parameter Values of the Antenna

Parameter	Description	Optimal Value
L	Length of the antenna	28 mm
W	Width of the antenna	21 mm
L_1	Length of the slot	15 mm
W_1	Width of the slot	16.8 mm
L_2	Length of the slot in the ground plane	1.8 mm
W_2	Width of the slot in the ground plane	16.2 mm
d	Feed gap distance	1.6 mm
H	Height of the patch	9.3 mm

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

The simulated VSWR, with and without slot in the ground plane is shown in Fig. 3 and it is inferred that the antenna with slot has a impedance bandwidth of 3-.1 GHz to 11.4 GHz for VSWR < 2 (S_{11} < -10 dB) with notch band from 5.1 GHz to 5.9 GHz.

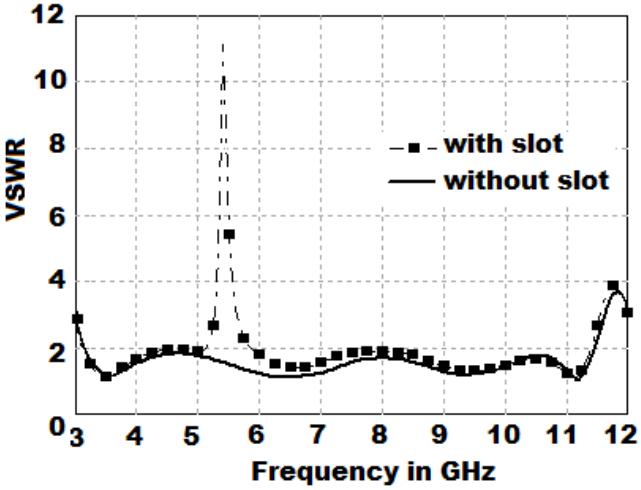
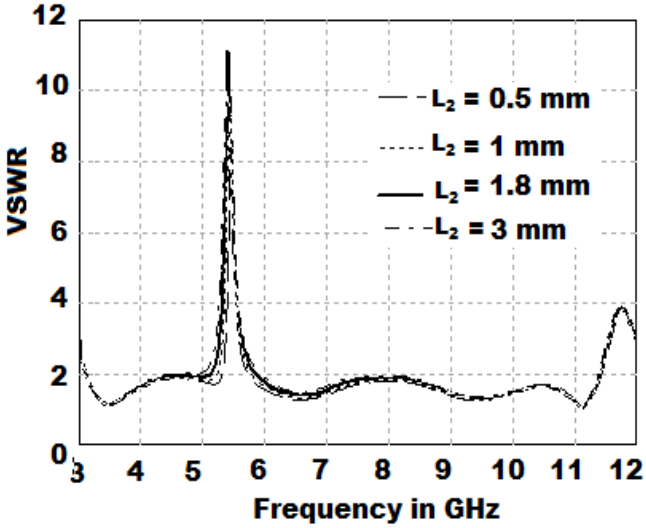


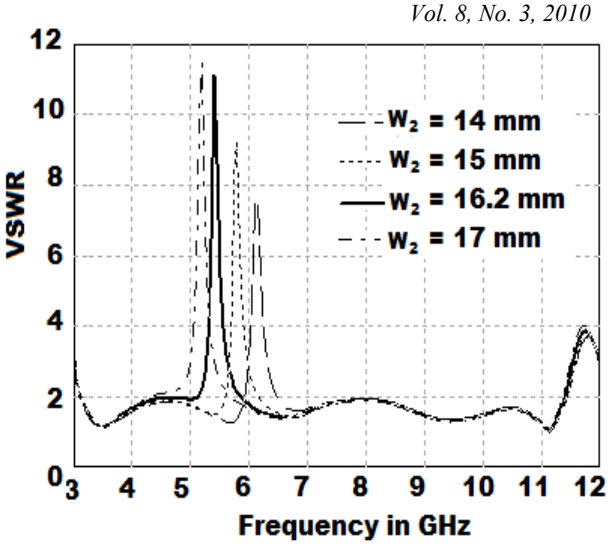
Fig. 3 Simulated VSWR with and without slot in the ground plane.

A. *Effect of Slot in the Ground Plane*

The slot is introduced in the ground plane as shown in Fig. 1 to obtain the notch in the specified band. Fig. 4a shows the VSWR variation for different slot lengths 'L_2' (for a particular width). It is observed that only a very small variation in the notch band for 28% of length variation with respect to the ground plane. Also it is noticed that the shift in notch band is directly proportional to slot length. Fig. 4b depicts VSWR variation for different widths'W_2' (L_2=1.8 mm optimal value). It is noticed that the notch frequency is shifted towards higher frequency when slot width is decreased and vice versa. The rate of notch frequency shift per unit width is 0.36 GHz. The optimal width value is 16.2 mm.



(a)

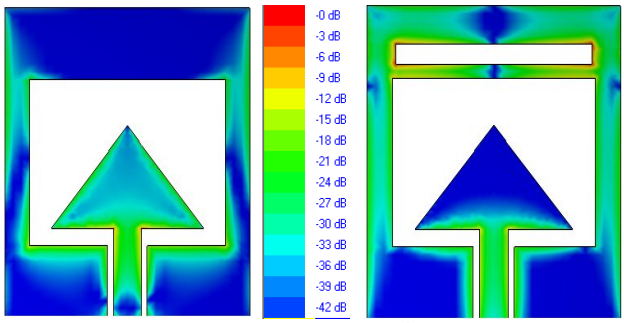


(b)

Fig.3 Simulated VSWR a) differnet slot lengths'L_2' b) different slot widths'W_2'

B. *Simulated Current Distribuition*

The simulated current distribution at frequency 5.5 GHz with and without slot in the ground plane is presented in Fig. 4. The current distribution of the proposed antenna is obtained by accounting the optimal design parameter values. In Fig. 4a the current distribution is mainly around the rectangular radiating slot and in the tuning stub. From the Fig. 4b, it is witnessed that the current distribution around the radiation slot is disturbed by the introduction of slot in the ground plane, which is responsible for the notch in the frequency band.



(a) (b)

Fig. 4 Current distributions at 5.5 GHz. a) without slot (b) with slot in the ground plane.

C. *Simulated Gain*

The computed gain of the UWB antenna with and without slot in the ground plane is compared in Fig. 5. It is observed that there is a gain variation at the notched frequency which is 0.5 dBi where as it is 3.5 dBi for the antenna without slot and the gain varies from 2 dBi to 6 dBi across the UWB spectrum.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

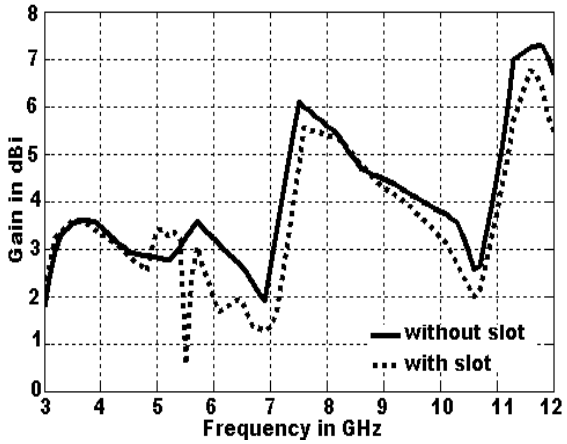


Fig. 5. Comparison of simulated gain responses.

D. Group delay perfromance

The group delay 'τ' of the antenna is calculated from the phase of the computed 'S_{21}' by using the following equation and plotted in Fig. 6,

$$\tau = - \frac{d\phi}{df} \qquad (2)$$

where 'ϕ' is phase of S_{21} in radians /sec and 'f' is frequency in GHz.

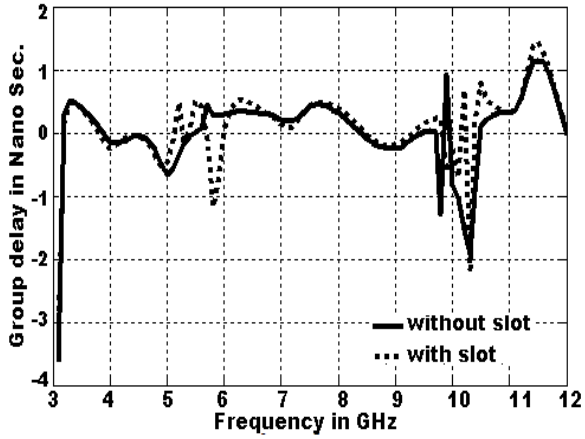


Fig. 6. Computed group delay.

From the Fig. 6, it is noticed that the variation in the group delay for the antenna with and without slot is around 2 ns for the frequency range from 3.1 GHz to 10.6 GHz, but there is small variation in the group delay at the notch band in the response for the antenna with slot which is acceptable one.

IV. Experimentall Results and Discussion

The prototype of the proposed antenna shown in Fig. 1 was fabricated for different parameters with their optimal values and tested, which is depicted in Fig. 7. Using Hewlett Packard Network Analyzer (HP8757D), the VSWR is measured and compared with the simulation result is shown in Fig. 8.

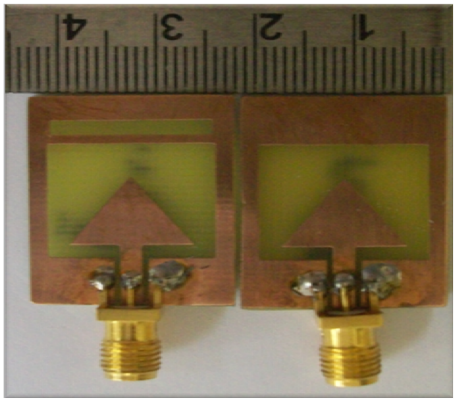


Fig. 7 Fabricated UWB slot antenna and its notch antenna.

There is a discrepancy between the measured and the simulated ones might be the effect of soldering the SMA connector or fabrication tolerance. The simulation result was obtained by assuming coplanar as input port, whereas practically SMA connector was used, the imperfect transition between SMA feed to coplanar may introduce losses [20] and also the capacitances can lead to shift in the frequency.

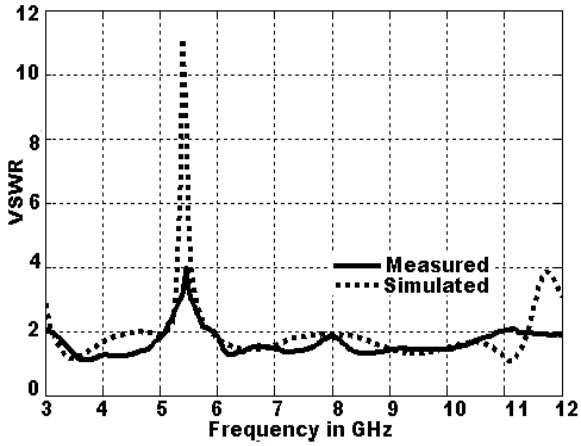


Fig. 8. Comparison of measured and simulated VSWR.

V. Time Domain Analysis

In ultra wideband systems, the information is transmitted using short pulses. Hence it is important to study the temporal behavior of the transmitted pulse. The communication system for UWB pulse transmission must provide as minimum as possible distortion, spreading and disturbance. The channel is assumed to be linear time invariant (LTI) system to verify the capability of the proposed antenna for transmission and reception of these narrow pulses. The transfer function of the entire system is computed using simulated value of 'S_{21}' parameter [21]. The received output pulse is obtained by taking the Inverse Fourier Transform (IFT) of the product of transfer function and spectrum of the test input pulse. While computing 'S_{21}', two identical antennas are placed face to face position at a distance of 70 cm that is greater than the far-field distance of the antenna.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

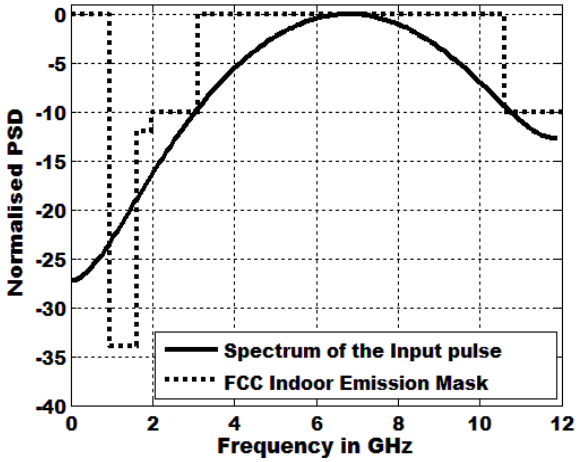


Fig. 9. FCC mask with the spectrum of the test input pulse.

The cosine modulated Gaussian pulse is considered for this analysis with centre frequency of 6.85 GHz and pulse width of 220 picoseconds, whose spectrum is shown in Fig. 9. It satisfies the requirement of FCC mask for UWB indoor emission. The comparison of input and output received pulse for the antenna with notch and without notch is shown in Fig. 10, which ensures the distortion less pulse transmission and also guarantees that the designed antenna is capable of transmitting and receiving short pulses. The ringing effect in the waveform may be due to the transmission properties of the system.

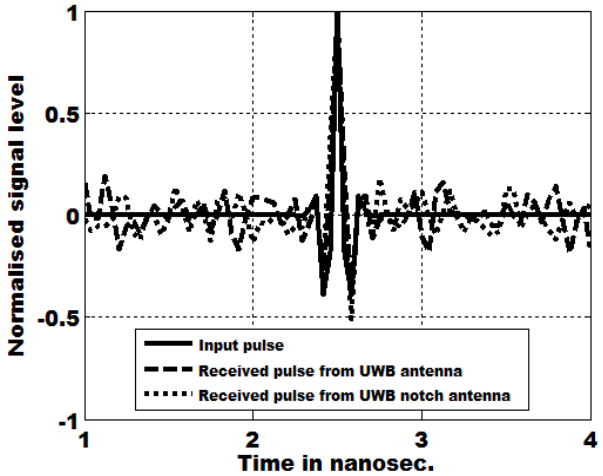


Fig. 10. Comparison of the received pulses from the UWB antenna without and with notch with input pulse.

VI. CONCLUSION

In this paper, a CPW fed UWB slot antenna with novel band notched design is proposed. The triangular stub is introduced at the anterior portion of the feed to enhance the coupling between the slot and feed. By embedding a slot in the ground plane, which is nearer to the radiating slot is responsible for the notch in the undesired band. With the above structural features the overall dimension of the proposed antenna comes around $28\times21\times1.6$ mm^3. The antenna covers the entire UWB system ranging from 3.1 GHz to 11.4 GHz with band notching between 5.1 GHz and 5.9 GHz. Time domain analysis of the proposed antenna ensures the suitability of the antenna for UWB applications. Hence, this antenna configuration would be quiet useful for UWB indoor applications with no interference from WLAN and HYPERLAN/2 systems when they coexist.

REFERENCES

[1] FCC NEWS(FCC 02-48), Feb. 14,2002. FCC News release, New public safety applications and broadband internet access among uses envisioned by FCC authorization of ultra wideband technology.

[2] M. Ghavami, L.B. Michael and R. Kohno " Ultra Wideband Signals and Systems in Communication Engineering", John Wiley and sons. Inc., NY, USA, 2004.

[3] K.L. Wong, " Compact and Broadband Microstrip Antenna", John Wiley and sons. Inc., NY, USA, 2001.

[4] I.-J. Yoon, H. Kim et al, "Ultra-wideband tapered slot antenna with band cutoff characteristic," *Electronics Letters*, vol. 41, no. 11,pp. 629-630, May 2005.

[5] A. M. Abbosh, M. E. Bialkowski et al, " A Planar UWB antenna With Signal Rejection Capability in the 4–6 GHz Band," *IEEE Microwave and Wireless Comp. Letters*, vol. 16, no. 5, pp. 278-280, May 2006.

[6] Keyvan Bahadori and Yahya Rahmat-Samii, "Miniaturized Elliptic-Card UWB antenna with WLAN Band Rejection for Wireless Communications," *IEEE Trans. Antennas and Propag*., vol.55, no.11. pp.3326-3332,Nov. 2007.

[7] Chong-Yu Hong, Ching-Wei Ling et al, " Design of a Planar Ultra- wideband Antenna With a New Band-Notch Structure ," *IEEE Trans. Antennas and Propag.* vol.55, no.12, pp.3391-3397,Dec. 2007.

[8] Q.-X. Chu and Y.-Y. Yang , " 3.5 / 5.5 GHz dual band-notch ultra-wide band antenna," *Electronics Letters,* vol. 44, no. 3, pp. 172-174, Jan. 2008.

[9] L. Wang, . Wu, X.-W. Shi et al , "Design Of A Novel monopole UWB Antenna with a notched ground," *PIER*, vol. 5,pp. 13–20, 2008.

[10] Jia-Yi Sze and Jen-Yi Shiu, " Design of Band-Notched Ultrawideband square aperture antenna with a hat-Shaped Back-Patch," *IEEE Trans. Antennas and Propag*., vol.56, no.10. pp.3391-3397,Oct. 2008.

[11] Yi-Cheng Lin and Kuan-Jung Hung, "Compact Ultrawideband Rectangular Aperture Antenna and Band-Notched Designs," *IEEE Trans. Antennas Propag.,* vol.54, no.11, pp.3075-3081, Nov.2006.

[12] Hyung Kuk Yoon, Yohan Lim et al, " UWB Wide Slot antenna with Band -notch Function," *Proc. IEEE Antennas and Propag. society Intl Symposium*, pp. 3059-3062, July 2006.

[13] Wen-jun Lui, Chong-hu Cheng and Hong-bo Zhu, "Improved Frequency Notched Ultra wideband Slot Antenna Using Square Ring Resonator," *IEEE Trans. Antennas Propag.,* vol.55, no.9, pp.2445-2450, Sept.2007.

[14] Yunlong Cai and Zhenghe Feng , " A UWB Antenna with novel L branches on ground for Band-Notching Application ," *Proc. Of IEEE Intl. Conference on Microwave and Millimeter wave Tec.*, vol. 4, pp.1654-1657, April 2008.

[15] The-Nan Chang and Min-Chi Wu , " Band-Notched Design for UWB Antennas , " *IEEE Antennas Wirel. Propag. Letters*, vol.7, pp.636- 639, 2008.

[16] C.- Y. Huang, S.- A. Huang and C.- F. Yang , " Band-notched ultra-wideband circular slot antenna with Inverted C-shaped parasitic strip,"*Electronics Letters* , vol. 44, no.15, pp. 891-892, July 2008.

[17] IE3D 14, Zeland Software, Ins., Fremont, USA.

[18] C. Balanis ,*Antenna Theory Analysis and Design*, 3rd edition, New York , Wiley, 2005

[19] J. William and R. Nakkeeran , "A CPW-fed wideband slot antenna with triangular patch," *Proc. of IEEE Intl. Conference on Computing,*

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Communication and Networking , Dec. 2008. DOI. 10.1109ICCCNET.2008.4787770.

[20] Kuang-ping ma, Yongxi Qian and Tatsuo Itoh, "Analysis and applications of a New CPW- slot line Transition," *IEEE Transactions on Microwave theory and Techniques*, vol. 47, pp. 426-432, April 1999.

[21] Stanislas Licul and William A Davis, " Ultra wide band (UWB) antenna measurements using vector Network analyzer," *IEEE Antennas and Propagation International Symposium,* pp. 1319-1322, June 2004.

AUTHORS PROFILE

J.William received his B.E. degree in Electronics and Communication from Bharathidasan University, Tamilnadu, India, and the M.Tech.degree in Communication Systems from National Institute of Technology (N.I.T), Trichy,India, in 1991 and 2006 respectively. He is currently working towards the Ph.D. degree at Pondicherry Engineering College, Pondicherry, India. He is a life member of ISTE and IE (I) and member of IEICE and EurApp. His current research interest is in the area of coplanar waveguide feed antennas and printed slot antennas for UWB.

R. Nakkeeran Received BSc. Degree in Science and B.E degree in Electronics and Communication Engineering from the Madras University in 1987 and 1991 respectively and M.E degree in Electronics and Communication Engineering (diversification in Optical Communication) from the Anna University in 1995. He received Ph.D degree from Pondicherry University in 2004. Since 1991, he has been working in the teaching profession. Presently, he is Assistant Professor in Pondicherry Engineering College. He is life member of IETE, ISTE, OSI and IE(I). Also he is member of OSA, SPIE and IEEE. He has published seventy five papers in National and International Conference Proceedings and Journals. He has co-authored a book, published by PHI. His areas of interest are Optical Communication, Networks, Antennas, Electromagnetic Fields and Wireless Communication.

A STUDY OF VARIOUS LOAD BALANCING TECHNIQUES IN INTERNET

M.Azath[1], Dr.R.S.D.Wahida banu[2],
[1]Research Scholar, Anna University, Coimbatore.
[1]mailmeazath@gmail.com
[2]Research Supervisor, Anna University, Coimbatore.
[2]drwahidabanu@gmail.com

Abstract

One of the most important applications of traffic engineering is load balancing. Successful implementation of load balancing depends on the underlying routing protocol that provides connectivity through the Internet by determining the routes used by traffic flows. But the load-balancing problem is not yet solved completely; new applications and architectures are required to meet the existing or incoming fastest Internet world. And, for greatest impact, these new capabilities must be delivered in toolkits that are robust, easy-to-use, and applicable to a wide range of applications. For balancing traffic in internet, packets should be reorder, reordering also having a problem for flows in internet. In Internet, unresponsive flows easily occupy the limited buffers, there by reducing the Quality of Service (QoS). In this paper, various techniques that are adopted for load balancing in Internet are analyzed.

Keywords:
Traffic engineering, Load Balancing, Internet Services, unresponsive flows, QOS, Buffer, Traffic splitting and Router.

1. Introduction

Traffic engineering refers to the performance optimization of operational networks. On one hand, traffic offered between origin and destination nodes loads the network and on the other hand, this traffic has to be carried in the network in such a way that performance objectives are fulfilled. In computer networking, load balancing is a technique to spread work between two or more computers, network links, CPUs, hard drives, or other resources, in order to get optimal resource utilization, throughput, or response time.

One of the most common applications of load balancing is to provide a single Internet service from multiple servers, sometimes known as a server farm. Commonly load balanced systems include popular web sites, large Internet Relay Chat networks, high bandwidth File Transfer Protocol sites, NNTP servers and DNS servers.

The idea of load balancing is to move traffic from congested links to other parts of the network in a well-controlled way. Traffic engineering seeks to effectively balance traffic load throughout existing networks, thus achieving QoS demands and minimizing typical costs of adding hardware and software implementations, common to network engineering. When dealing with real-world cases of load balancing, both network and traffic engineering is general purpose tools used throughout all steps of an implementation [12].

2. Motivation

Load balancers are an integral part of today's Web infrastructure. They're also complex and under-documented pieces of hardware.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Today's Web sites are complex beasts. Every component must work together to create a site that is greater than the sum of its parts.

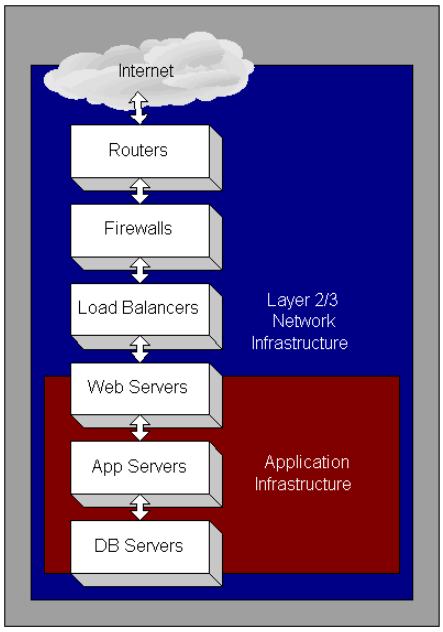


Figure 1: Traffic flow for a load balancer

The Internet is connected to the routers, which pass traffic through a firewall to the load balancers, which distribute the traffic to the Web servers, which pass information to the application server bone, and the application server bone is connected to the database server bone. We get the picture. If one component or piece of the process fails, it can take down the entire site.

Load balancers are also in the direct path of all traffic to a particular Web site. By looking at Figure 2 below, we can see that if the load balancer stops working, the entire site stops working. This critical position in the infrastructure can make it appear as though the load balancer is the problem, even in cases where it is not (such as a firewall issue, a back-end database problem, someone tripping over a cable, etc.). Unlike a broken or malfunctioning Web server, a miss configured or malfunctioning load balancer will result in a dead-to-the-world site. This is why a firewall is often a suspect, too, but to a lesser degree since it is generally a simpler device than load balancers [1].

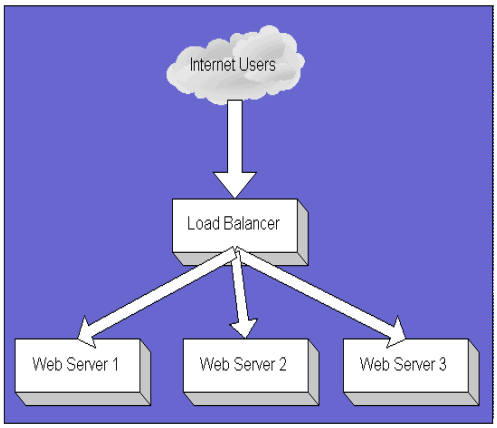


Figure 2: Load Balancer implementation

3. Needs of load balancing

Load balancing is common in ISP networks. If the traffic demands are known, the load balancing can be formulated as an optimization problem. However, knowledge of traffic demands is often lacking. Recent trends in load balancing points toward dynamic protocols [13]. These protocols map the traffic of router onto multiple paths and adapt the share of each path in real-time to avoid hot-spots and cope with failures. Dynamic load balancing needs schemes that split traffic across multiple paths at a fine granularity.

Since the power of any server is finite, a web application must be able to run on multiple servers to accept an ever increasing number of users. This is called scaling. Scalability is not really a problem for intranet applications since the number of users has little chances to increase.

However, on internet portals, the load continuously increases with the availability of broadband Internet accesses. The site's maintainer has to find ways to spread the

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

load on several servers, either via internal mechanisms included in the application server, via external components, or via architectural redesign.

Further complication increases due to various factors such as:

- Sizes of objects might not be the same.
- Object IDs might not be chosen at random.
- Heterogeneity in the capabilities of nodes.

4. Types of load balancing and its approaches

Load balancing algorithms can be classified into three main classes: static algorithms, dynamic algorithms, and adaptive algorithms. Static algorithms decide how to distribute the workload according a prior knowledge of the problem and the system characteristics. Dynamic algorithms use state information to make decisions during program execution. Finally, Adaptive algorithms are a special case of dynamic algorithms. They dynamically change its parameters in order to adapt its behavior to the load balancing requirements [10].

As though there are three main algorithms for load balancing, there are different load balancing approaches available. They are:

1. Client-side load balancing: Client-side load balancing is not a normal practice, but it is indeed possible. For some times, Netscape incorporated a simple balancing algorithm in their Navigator browser, making it choose a random Netscape web-server when visiting www.netscape.com. [2].
2. Core Internet routing uses protocols and agreements that allow for automated load balancing and fail-over mechanisms. These are commonly based on the Border Gateway Protocol, used for data-exchange between large Internet operators.
3. DNS-based load balancing is a popular way of distributing traffic amongst a set of Internet addresses by returning a list of active addresses to the requesting client. These addresses can point to a set of servers or even a set of geographically separate sites.
4. Sites can connect to the net through several links, a practice known as multi-homing. This enables both incoming and outgoing load balancing, in addition to the increased redundancy.
5. Dispatcher-based load balancing is used within a site to balance load between a set of real servers. Generally, the dispatcher assumes a virtual address for a service and receives requests which it then redirects to an appropriate server based on given criteria.
6. The real servers can again operate some form of balancing mechanism to decide whether to handle the request or redirect it to a more suitable server or site.
7. Content servers could access back-end servers – typically running databases and low-level services – in a load-balanced fashion.
8. Back-end servers could also incorporate balancing amongst themselves to avoid over-utilization.

When talking about varying levels of load balancing, it is fair to identify the level to be proportional to the distance from the content served – long distance equals high level, and vice versa. In an informal manner, we can designate steps 1through 4 in the figure as high-levels, and steps 5 through 8 as low-levels of load balancing [12].

5. Adopted techniques

There are various papers which explain about the load balancing in Internet. Each of

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

them is routed in different directions and with various considerations.

Router mechanisms designed to achieve fair bandwidth allocations, like Fair Queueing, have many desirable properties for congestion control in the Internet. In [6], architecture is proposed that significantly reduces this implementation complexity yet still achieves approximately fair bandwidth allocations. This architecture has two key aspects:
1. To avoid maintaining per flow state at each router, we use a distributed algorithm in which only edge routers maintain per flow state, while core (non-edge) routers do not maintain per flow state but instead utilize the per-flow information carried via a label in each packet's header. This label contains an estimate of the flow's rate; it is initialized by the edge router based on per-flow information, and then updated at each router along the path based only on aggregate information at that router.
2. To avoid per flow buffering and scheduling, as required by Fair Queueing, we use FIFO queueing with probabilistic dropping on input. The probability of dropping a packet as it arrives to the queue is a function of the rate estimate carried in the label and of the fair share rate at that router, which is estimated based on measurements of the aggregate traffic. The limitation here is in the destination, packet reorganizing should be there.

In [18], they evaluated 5 direct hashing methods and one table-based hashing method. While hashing schemes for load balancing have been proposed in the past, this is the first comprehensive study of performance using real traffic traces. They find that hashing using 16-bit CRC over TCP five tuples gives excellent load balancing performance. Load-adaptive table-based hashing uses the exclusive-OR of the source and destination IP addresses achieve comparable performance to 16-bit CRC. As though hashing methods provide best performance, new hash based algorithms are needed that have less computational complexity.

In [7], they proposed a per-class queue management and adaptive packet drop mechanism in the routers for Internet congestion control. An active queue management is modeled as an optimization problem and the proposed mechanism provides congestion control and fairness for different types of traffic flows. An optimal packet drop rate is obtained to maintain a relatively small queue occupancy, which provides a less queue delay delivery of packets. Moreover, the queue occupancy and the packet drop rates obtained are both upper bounded, which is meaningful for providing the class-based guaranteed delay services for real-time multimedia applications. They model the general AQM as an optimization problem, and try to obtain a minimal packet drop rate that results in low queue occupancy. Compared with RED that controls the average queuing delay in the router, the per-class queue management and optimal packet drop mechanism can obtain the minimal queuing delay and hence the end-to-end delay. The major drawback in this paper is that the packet drop rate is minimal and the resulting queue occupancy is also kept minimal.

In [19], a novel packet scheduler is proposed and is called Stratified Round Robin, which has low complexity, and is amenable to a simple hardware implementation. In particular, it provides a single packet delay bound that is independent of the number of flows. This property is unique to Stratified Round Robin among all other schedulers of comparable complexity.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

An important component of the many QoS architectures proposed is the packet scheduling algorithm used by routers in the network. The packet scheduler determines the order in which packets of various independent flows are forwarded on a shared output link. One of the simplest algorithms is First Come First Served (FCFS), in which the order of arrival of packets also determines the order in which they are forwarded over the output link. FCFS clearly cannot enforce QoS guarantees, as it allows rogue flows to capture an arbitrary fraction of the output bandwidth.

Stratified Round Robin operates as a two-step scheduler:

1. The first step uses the flow class mechanism to assign slots to each flow fi in proportion to its approximate weight as defined by the flow class Fk to which it belongs.
2. The second step uses the weight-proportional credit mechanism to ensure that each flow fi receives service in proportion to its actual weight wi.

The advantage of this approach is that it considerably simplifies the scheduling decision to be made. But we are having the difficulty of considerable packet reorganizing in the destination.

In this paper [11], based on measurements of Internet traffic, they examined the sources of load imbalance in hash-based scheduling schemes. They proved that under certain Zipf-like flow-size distributions, hashing alone is not able to balance workload. They introduced a new metric to quantify the effects of adaptive load balancing on overall forwarding performance. To achieve both load balancing and efficient system resource utilization, they proposed a scheduling scheme that classifies Internet flows into two categories: the aggressive and the normal, and applies different scheduling policies to the two classes of flows. They have stated that their work is unique in exploiting flow-level Internet traffic characteristics.

In [13], a new mechanism called Flow-let Aware Routing Engine (FLARE) is proposed in which a new traffic splitting algorithm that operates on bursts of packets, carefully chosen to avoid reordering. Using a combination of analysis and trace-driven simulations, it is shown that FLARE attains accuracy and responsiveness comparable to packet switching without reordering packets. FLARE is simple and can be implemented with a few KB of router state. Highly accurate traffic splitting can be implemented with little to no impact on TCP packet reordering and with negligible state overhead. Owlets can be used to make load balancing more responsive, and thus help enable a new generation of real-time adaptive traffic engineering.

Static load balancing is presented in [3]. The static problem is possible to be formulated and solved only if we have precise information on all the traffic demands. However, such information is not available or traffic condition change unexpectedly. In dynamic load balancing, dynamically changing network status information are utilized [5],[8],[14],[4],[15].

Traffic Engineering (TE) of dynamic methodologies is classified into two basic types: time-dependent and state-dependent. In time-dependent TE, traffic control algorithms are used to optimize network resource utilization in response to long time scale traffic variations. In state-dependent TE, traffic control algorithms adapt to relatively fast network state changes. State-dependent load balancing is a key technique for improving the performance and

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

scalability of the Internet. The fundamental problem in dynamic load balancing in distributed nodes involves moving load between nodes. Each node can transfer load to at most one neighbor also, any amount of load can be moved along a communication link between two nodes in one step. The dynamic load balancing is formed with incomplete information. More precisely, it is assumed that the traffic demands are unknown but the link loads are periodically measured using a measurement system as in [9], [16] [17].

Floyd *et al.* [20] study congestion collapse from undelivered packets. This situation arises when bandwidth is continuously consumed by packets at the upstream that are dropped at the downstream. Several ways to detect unresponsive flows are presented. It is suggested that routers can monitor flows to detect whether flow is responsive to congestion or not. If a flow is not responsive to congestion, it can be penalized by discarding packets to a higher rate at the router. According to the authors there are some limitations of these tests to identify non-"TCP-friendly flow". It does not help to save bandwidth at the upstream if the flow sees the congestion at the downstream because this solution does not propagate the congestion information from downstream to upstream [21].

In [22], architecture which contains an adaptive packet scheduler with a bursty traffic splitting algorithm is proposed for better load balancing. The scheduler has a classifier which classifies the flows into aggressive and normal flows. Aggressive flows are treated as high priority flows. Based on the buffer occupancy threshold, a trigger handler checks for load un-balance of the network and automatically triggers the load adapter. The load adapter reroutes the high-priority aggressive flows into the least loaded best path, using the bursty traffic splitter algorithm. As though it is adaptive algorithm for load balancing, it does not satisfy the needs when we are coming across the unresponsive flows.

6. Conclusion

Static load balancing is suffering from lack of all the available information while we perform for load balancing. Similarly dynamic load balancing needs schemes that split traffic across multiple paths at a fine granularity. In adaptive load balancing the performance of the network can be improved by parameter adjustments of the routes, traffic splitting, and scheduling. Thus in this current fast Internet scenario, a new adaptive load balancing algorithm is needed. The adaptive load balancing is also not sufficient for some situations like unresponsive flows of traffics. Hence there is a need that for load balancing algorithm with the additional feature of detecting unresponsive flows of the incoming traffic in Internet.

References

[1] http://www.oreillynet.com/pub/a/oreilly/networking/news/slb_0301.html, Feb 2010.

[2] Dan Mosedale, William Foss, and Rob McCool, "Lessons learned administering netscape's internet site." *IEEE Internet Computing*, 1(2):28– 35, 1997.

[3] Grenville Armitage, 2000. MPLS the magic behind the myths. IEEE Commun. Mag., 38: 124-131.

[4] Dinan, E., D. Awduche and B. Jabbari, "Analytical framework for dynamic traffic partitioning in MPLS networks", IEEE International Conference on Communications, (ICC'00), 18-22 June 2000, New Orleans, Volume-3, pp: 1604-

1608.

[5] Elwalid, A., C. Jin, S. Low and I.Widjaja, MATE: MPLS adaptive traffic engineering. IEEE Infocom., Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies, Volume-3, April 22-26 2001, Anchorage, Alaska, USA. ISBN 0-7803-7016-3, 1300-1309.

[6] Core-Stateless Fair Queuing: A Scalable architecture to approximate fair bandwidth allocations in high speed networks, Ion Stoica, Scott Shenker, Hui Zhang, 2003.

[7] Mei-Ling Shyu, Shu-Ching Chen, Hongli Luo, "Per-class queue management and adaptive packet drop mechanism for multimedia networking", Proceedings of the 2003 International Conference on Multimedia and Expo - Volume 3 (ICME '03) - Volume 03.

[8] Song, J., S. Kim, M. Lee, H. Lee and T. Suda, "Adaptive load distribution over multipath in MPLS networks". IEEE International Conference on Communications (ICC'03), Anchorage, Alaska, Date: 11-15 May 2003 pp: 233-237.

[9] Butenweg, S., "Two distributed reactive MPLS traffic engineering mechanism for throughput opt'imization in Best effort MPLS networks, In the Eighth IEEE International Symposium on Computers and Communications (ISCC'03). 379-384 vol.1, 30th June-3rd July 2003.

[10] Mohammed Aldasht, Julio Ortega, Carlos G. Puntonet; Antonio F. Diaz, "A Genetic Exploration of Dynamic Load Balancing Algorithms", IEEE 2004.

[11] Shi, W. Macgregor, M.H. Gburzynski, P. "Load Balancing For Parallel Forwarding" Networking, Ieee/Acm Transactions On, Aug. 2005.

[12] Sven Ingebrigt Ulland, "High level load balancing for web services",University of Oslo, 20th May 2006.

[13] Srikanth Kandula, Dina Katabi, Shantanu Sinha, Arthur Berger "Dynamic Load Balancing Without Packet Reordering" Acm Sigcomm Computer Communication Review ,Volume 37 , Issue 2 (April 2007).

[14] Murugesan, G. and A.M. Natarajan, "Adaptive granularity algorithm for effective distributed load balancing and implementation in multiprotocol label switching networks". IEEE, International Conference on Advanced Computing and Communication (ADCOM'07) 18-21 December 2007, pp: 626-631.

[15] Dengyin Zhang, Zhiyun Tang and Ruchuan Wang, "Automatic traffic balance algorithm based on traffic engineering". J. Network Syst. Manage., 14: 317-325.

[16] Ashwin Sridharan, Roch Guerin and Christophe Diot, "Achieving near-optimal traffic engineering solution for current OSPF/IS-IS networks". In the IEEE/ACM Trans. Network., 13: 234-247.

[17] G. Murugesan, A.M. Natarajan and C. Venkatesh, "Enhanced Variable Splitting Ratio Algorithm for Effective Load Balancing in MPLS Networks" Journal of Computer Science 4 (3): 232-238, 2008 ISSN 1549-3636 2008 Science Publications.

[18] Zhiruo Cao, Zheng Wang, Ellen Zegura, "Hashing-based traffic splitting algorithms for Internet load balancing", Georgia Institute of Technology 1999.

[19] S. Ramabhadran and J. Pasquale, "Stratified Round Robin: A Low

Complexity Packet Scheduler with Bandwidth Fairness and Bounded Delay," Proc. Acm Communications Architectures and Protocols Conf. (Sigcomm), Karlsruhe, Germany, Pp. 239-249, Aug. 2003.

[20] S. Floyd and K. Fall. Promoting the use of end-to-end congestion control in the Internet. *IEEE/ACM Transactions on Networking*, Aug. 1999.

[21] Ahsan Habib, Bharat Bhargava, “Network Tomography-based Unresponsive Flow Detection and Control”, Department of Computer science, Purdue University, IN 47907-1398.

[22] M.Azath, Dr.R.S.D.Wahida banu, “Load balancing in Internet Using Adaptive Packet Scheduling and Bursty Traffic Splitting”, International Journal of Computer Science and Network Security, Vol.8, No.10, Oct 2008.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Laboratory Study of Leakage Current and Measurement of ESDD of Equivalent Insulator Flat Model under Various Polluted Conditions

N. Narmadhai
Senior Lecturer, Dept of EEE
Government College of Technology
Coimbatore, India
narmadhai@gct.ac.in

S. Suresh
PG Scholar, Dept of EEE
Government College of Technology
Coimbatore, India
suresh.sundararaju@gmail.com

Dr.A.Ebenezer Jeyakumar
Director (Academics)
SNR Sons Charitable Trust, SREC
Coimbatore, India
ebeyjkumar@rediffmail.com

***Abstract*—The phenomenon of flashover in polluted insulators has been continued by the study of the characteristics of contaminating layers deposited on the surface of insulators in high voltage laboratories. This paper proposed the Equivalent insulator flat plate model for studying the flashover phenomena due to pollution under wet conditions even at low voltage. Laboratory based tests were carried out on the model under AC voltage at different pollution levels. Different concentrations of salt solution has been prepared using sodium chloride, Kaolin and distilled water representing the various contaminations. Leakage current during the experimental studies were measured. A conductivity measuring instrument (EQ-660 A) is used to measure the conductivity of the salt-solution. Salinity and Equivalent salt Deposit Density (ESDD) were calculated. Test results in terms of conductivity and ESDD are plotted against salt concentration and the relationship between the conductivity and ESDD is examined.**

Reported results of this preliminary study on the insulator model simulates the distinctive stages of development of flashover due to the pollution and it could be easily identified from the contamination level of ESDD and from the magnitude of leakage current.

Keywords-Conductivity, ESDD, Flashover, Insulator model, Leakage Current, Salt solution.)

I. INTRODUCTION

Insulators used in outdoor electric power transmission lines are exposed to outdoor environmental contaminations. Contamination on outdoor insulators enhances the chances of flashover. Depending on the nature and duration of exposure, deposits of wind-carried industrial, sea and dust contaminants build up on the insulator surface as a dry layer. The leakage current path through a layer of dry contaminants on an insulator surface is capacitive wherein the current amplitude is small and sinusoidal. The dry contaminant layer becomes conductive when exposed to light rain or morning dews. As wetting progresses, the leakage current path changes from capacitive to resistive with simultaneous increase in current amplitudes. The increase in leakage current dries the conducting layer and forms the dry bands around the areas with high current density. These dry bands interrupt the current flow and most of the applied voltages are impressed across these narrow dry bands. If the dry bands cannot withstand the voltage, localized arcing develops and the dry bands will be spanned by discharges. The arcs merge together and form a single arc, which triggers the surface flashover [1].

The contamination severity determines the frequency and intensity of arcing and, thus the probability of flashover. In favourable conditions when the level of contamination is low, layer resistance is high and arcing continues until the sun or wind dries the layer and stops the arcing. Continuous arcing is harmless for ceramic insulators. The mechanism described above shows that heavy contamination and wetting may cause insulator flashover and service interruptions. Contamination in dry conditions is harmless. B. F. Hampton investigated the voltage distribution along the wet, polluted surface of a flat insulating strip and the method of dry band formation, with subsequent growth of discharges on the polluted surface [2]. Verma measured the peak leakage current and correlated the current with the flashover voltage. He suggested that the flashover is imminent if the leakage current peak exceeds 100mA [3]. Karady observes the same [4].

In practice, there are various contaminant types that settle on outdoor insulators. These contaminants can be classified as soluble and insoluble. Insulators located near coastal regions are typically contaminated by soluble contaminants, especially salt (or sodium chloride).Insulators located near cement or paper industries are typically contaminated by non-soluble contaminants such as calcium chloride, carbon and cement dust. Irrespective of the type of contaminant, flashover can occur as long as the salts in the contaminant are soluble enough to form a conducting layer on the insulator's surface. In order to quantify the contaminants on the surface of the insulators, the soluble contaminants are expressed in terms of Equivalent Salt Deposit Density (ESDD), which correlates to mg of NaCl per unit surface area. Non-soluble contaminants are expressed in terms of Non-Soluble Deposit Density (NSDD), which correlates to mg of kaolin per unit surface area.

Many researchers studied that the leakage current due to the contamination level is the main cause for flashover. M.A.M.Piah and Ahmed Darus [5] modelled the leakage

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

current in terms of conductivity and other environmental parameters. I.R. Vazquez et al [6] have tested the non ceramic insulators by non standard method where they expressed the contamination level by ESDD. F. V. Topalis et al [7] studied the critical flashover voltage of the insulator with the variation of surface conductivity and ESDD. In Guan Zhicheng et al [11] the maximum value of leakage current has the definite relationship with the flashover voltage of the polluted insulator used to express the pollution degree of insulator.

In this paper, flat plate model for studying the insulator flashover phenomena has been presented. Experimental tests have been conducted on the model under various polluted conditions. The study on the leakage current is important in order to diagnose the insulator condition by monitoring of the leakage current. Here both ESDD and surface conductivity for various quantities of contaminated salt deposits are measured using the IEC standard [8].

II. Flat Plate Model and Experimental Setup

A. Equivalent Insulator Model

The simplified geometrical models equivalent to actual insulator are being widely used for the purpose of flashover analysis. Among these models, the basic flat trough model has merited extensive attention in the context of pollution flashover. So the proposed model, equivalent to standard disc insulator made of an insulating glass material with two copper terminals, one on cap and another at the pin. A simplified plan of the insulator model is shown in Fig. 1

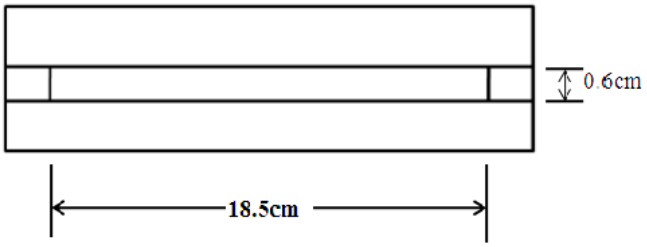


Fig. 1. Flat Trough Model

B. Experimental Setup

A proposed equivalent insulator trough model [1] of dimension18.5x 0.6x 0.2cm is used for the contamination flashover experiments. The principal application of this equivalent model would be to help simulate as much as possible the practical conditions of high voltage insulators in the application of low voltage itself. In artificial testing, a contaminant is usually substituted by a dissolved mixture of an inert binder-Kaolin and NaCl salt. The inert binder is supposedly non-conducting and the quantity of salt represents the level of contamination. Contamination salt solution was prepared for various NaCl values of 15g, 20g, 25g, and 30g. The mixture, usually dissolved in distilled water is known as slurry which is thoroughly mixed as per IEC standard [8]. Before coating, the trough is initially washed and wiped clean and dry. The experimental setup to measure the leakage current is shown in Fig 2. The slurry is poured so that it rolls off uniformly in the trough.

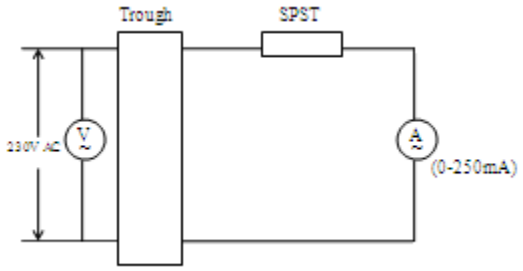


Fig. 2. Experimental Setup

A test voltage of 230V, 50 Hz was applied across the terminals and the leakage current is monitored through the suitable measuring meter from the instant of application of voltage till the formation of dry band. The dry band was precisely located on the model. Its shape, contour of growth and locations were physically measured. The test results either in a flashover or a withstand.The conductivity and ESDD has also been calculated from the deposited contaminations. The contamination can thus be classified as light, medium or heavy according to the IEC standard [9].

III. ESDD Calculation

In any insulator severity of pollution is characterized by the Equivalent Salt Deposit Density (ESDD). The procedure for calculating the ESDD [8] is as follows: After the test has been completed, the deposits were collected by a small brush from the contaminated plate and mixed with 1litre of distilled water to get the solution for specific area of the glass plate. This process is repeated for the other samples of salt solutions. The conductivity of each collected salt solution is measured using a conductivity meter which is initially calibrated using 0.1N KCl solution. At the same time temperature is also recorded. The conductivities at different temperature are converted to 20° temperatures by the expression [8] as,

$$\sigma_{20} = \sigma_{\theta}[1 - b(\theta - 20)] \quad (1)$$

Where,
θ is the solution temperature, °C
σ_{θ} is the volume conductivity at a temperature θ°C (S/m)
σ_{20} is the volume conductivity at a temperature 20°C (S/m)
b is the factor depending on the temperature θ as given in Table I.

Table I values of b at different temperatures

θ°C	b
5	0.03156
10	0.02817
20	0.02277
30	0.01905

The salinity S_a of the solution is determined by the following expression [8] as,

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

$$S_a = (5.7\sigma_{20})^{1.03} \quad (2)$$

Finally, the equivalent salt deposit density can be determined by the following expression [8] as,

$$ESDD = \frac{S_a \times V}{A} \quad (3)$$

Where,
V is the volume of the solution, cm^3
A is the area of the cleaned surface, cm^2

According to IEC 60815, pollution site severity classification are shown in Table II [9]

Table II Pollution Site severity (IEEE definitions)

Description	ESDD
Very Light	0-0.03
Light	0.03-0.06
Medium	0.06-0.1
Heavy	>0.1

IV. Artificial Contamination Test

A. Light Contamination

Insulators are mostly affected by flashover due to the deposition of NaCl salt particles. Therefore the equivalent model was uniformly sprayed with the slurry solution consisting of 15g NaCl, 40g Kaolin and 1 litre of distilled water. Leakage current started to flow on the surface of insulator due to the pollution. It is observed that there is an increase in leakage current magnitude when compared with clean surface condition, which is mainly because of increase in surface conductivity. Dry bands have started to form on the polluted surface after reaching the maximum leakage current of 42mA. The complete dried condition of pollution layer with complete dry band was physically seen after 60 minutes approximately. No flashover could be seen and the test results in a withstand.

Similarly for 20g NaCl, the leakage current increased to 90mA and the time taken for the formation of dry band is reduced compared to 15g.The test results in a Withstand.

B. Medium Contamination

Experiments were repeated for 25g NaCl. It is noticed that the magnitude of maximum leakage current increased to 130mA. It is because that the current magnitude depends on the level of contamination and the amount of moisture on the insulator surface. The test results in withstand but the field exceeds the withstand capability and it initiates the arc discharge.

C. Heavy Contamination

Finally the insulator model is contaminated by 30g NaCl solution and experiments were repeated in a similar way. Due to the high contamination of NaCl the magnitude of leakage current goes upto 220mA. The high magnitude of leakage current caused the heating effect which leads to rapid dry band formation and partial discharges across these dry bands. Due to the higher resistance at pin and cap end the heat dissipated in that location may be greater and therefore moisture dried rapidly. After 10 minutes dry bands could not sustain the applied voltage cause the scintillations to occur which ultimately leads to flashover. The scintillations have been physically seen and captured using high speed camera.

V. Results and Discussions

A. ESDD

The correction conductivity, salinity and ESDD have been calculated for various tests as per IEC-507. The measured Salinity, ESDD and conductivities for 15g, 20g, 25g and 30g of NaCl salt are shown in Table III.

Table III Values of conductivity, Salinity & ESDD using salt solutions

NaCl	θ	σ_θ	σ_{20}	S_a	ESDD
15	28	0.114	0.0966	0.000439	0.0396
20	28	0.165	0.1398	0.000643	0.0579
25	28.5	0.205	0.1718	0.000795	0.072
30	28	0.315	0.2669	0.001252	0.1128

The correction conductivity and ESDD obtained from the present measurements by varying the amount of salt concentration is presented in Fig. 3.

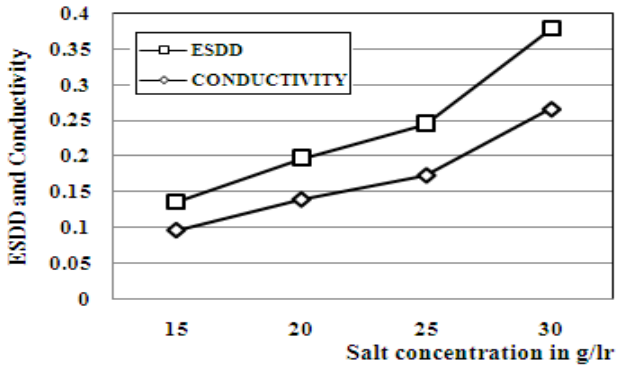


Fig. 3. Variation of ESDD and Conductivity with salt concentration

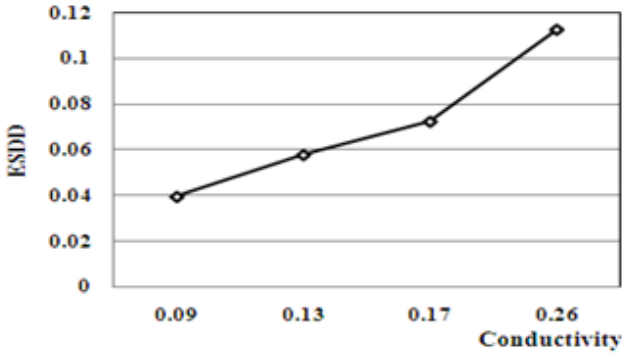


Fig. 4. Variation of ESDD with conductivity

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

In Fig. 3, for the salt concentration from 15g to 30g, the variation of conductivity and ESDD is almost linear. The values of ESDD are also plotted versus the conductivity which is shown in Fig. 4. It is observed that the relationship between the conductivity and ESDD is linear [10].

B. Leakage Current

Table IV Experimental test result

Applied voltage(V)	NaCl Qty(g)	Leakage Current(mA)		Status of Flashover
		Initial Value	Maximum Value	
230	15	33	42	Withstand
230	20	58	96	Withstand
230	25	90	130	Withstand
230	30	106	220	Flashover

Table V Test results for various pollution levels

NaCl Qty(g)	Leakage Current(mA)		ESDD (mg/cm^2)	Pollution Level
	Initial Value	Maximum Value		
15	33	42	0.0396	Light
20	58	96	0.0579	Light
25	90	130	0.0725	Medium
30	106	220	0.1128	Heavy

The following points were observed from the results shown above in Table IV and V.

i. For light contamination of 15g NaCl, the maximum leakage current measured is 42 mA. The calculated ESDD from the contamination deposit is 0.0396 which indicates the light degree of pollution.

ii. For 20g NaCl, maximum leakage current is 96mA which results in withstand and the corresponding ESDD of 0.0579 also indicates the low level of pollution.

iii. For 25g the leakage current reaches a peak of 130mA the field exceeds the withstand capability, initiates an arc discharge and extends several arcing which is actually preceded before the flashover. ESDD value is 0.0725 shows the medium level of pollution.

iv. For 30g the leakage current measured is 106mA and reached a peak of 220mA. It results in flash over after 10 minutes of wetting. ESDD shows the heavy pollution level for 0.1128.

v. The leakage current shows that the pollution severity can be correlated with the surface conductivity and ESDD. When these quantities reach the critical value, the flashover is imminent [3] [11].

vi. Using the proposed model, the leakage current were measured and found similar [3] [4] from experiment. These revealed the equivalent model with 230V supply could be used for flashover prediction and analysis in the place of the standard high voltage insulator.

VI. CONCLUSION

To simplify the mathematical analysis, tests were done on a flat trough model of simple geometry. The maximum leakage current and the conductivity of the contaminant solution on the surface of the model and the corresponding ESDD of various degree of pollution were also determined. The laboratory model test results either in a flashover or a withstand. The result also showed that the leakage current strongly correlated with insulator polluted level and to assess the condition of the insulation system which indicates the occurrence of a flashover situation and the need for cleaning of the insulators when it exceeds the medium level of ESDD.

Even though the model presented above still needs modifications, the extent of study of leakage current magnitude for the analytical case is satisfactory. At low voltage, the model tests can be standardized easily; its use in studying contamination flashover should be encouraged. In this way, fairly accurate results could be obtained eliminating the need of site testing.

REFERENCES

[1] Ebenezer Jeyakumar, *"Development of verisimilar juxtaposition model and study of physical phenomena on polluted insulators,"* PhD Dissertation, Department of Electrical Engineering, Anna University Madras, India, June 1991.

[2] B. F. Hampton, *"Flashover mechanism of polluted insulation,"* Proc.IEE, vol.111, pp.985-990, 1964.

[3] M.P.Verma, *"Die quantitative erfassung von fremdschichichteintlusse,"* ETZ-A97, pp. 281-285, 1976.

[4] George Karady, Felix Amarh, Raji Sundararajan, *"Dynamic modeling of ac insulator flashover characteristics,"* 11th International Symposium of High Voltage Engineering, Vol. 4, pp. 107-110, London, England, August 1999.

[5] M.A.M. Piah, Ahmad Darus, *"Modeling leakage current and electric field behavior of wet contaminated insulators,"* Power Engineering Letters, IEEE Transactions on Power Delivery 19 (1) 432–433, January 2004.

[6] I.R. Vazquez, G.M. Tena, R.H. Corona, *"Nonstandard method for accelerated aging tests of non ceramic insulators,"* IEE Proceedings – Generation Transmission and Distribution 149 (4), 439–445, July 2002.

[7] F. V. Topalis, I. F. Gonos and I. A. Stathopulos, *"Dielectric behavior of polluted porcelain insulators,"* IEE Proc.-Gener. Transm. Distrib., Vol 148, No. 4, pp. 269-274, July 2001.

[8] IEC 60507, *"Artificial pollution tests on high voltage insulators to be used in ac system,"* Switzerland,1991.

[9] IEC60815, *"Guide for the selection of insulators in respect of polluted conditions,"* 1986.

[10] M. A. Salam, N Mohammad, Zia Nadir, Ali Al Maqrashi, A Al Kaf, *"Measurement of conductivity and equivalent salt deposit density of*

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

contaminated glass plate," IEEE conference TENCON, Vol.3 pp 268-270, 2004.

[11] Guan Zhicheng, Cui Guoshun, *"A study on the leakage current along the surface of polluted insulator,"* Properties and applications of dielectric materials, Proceedings of the 4th international conference on volume 2, 3-8 July, pp.495-498,1994

SSL/TLS Web Server Load Optimization using Adaptive SSL with Session Handling Mechanism

R.K.Pateriya[#], J.L.Rana[#], S.C. Shrivastava[#]
[#]Department of Computer Science & Engineering and Information Technology
Maulana Azad National Institute of Technology, Bhopal, India
Emails: pateriyark@gmail.com , jl_rana@yahoo.com , scs_manit@yahoo.com

Abstract — Secure Sockets Layer (SSL) is the world standard for web security. SSL provide authentication ,data integrity and ensure message confidentiality using cryptography. This paper proposes an approach for load management by applying Adaptive SSL (ASSL) policy with session handling mechanism for enhancement of the security and performance of the server . ASSL policy negotiate session security at runtime by adapting more secured and comparatively costly cryptographic algorithm at runtime if load is under safe limit otherwise change to less secure algorithm .Session handling mechanism limit the active session running on the server .This self-adaptive security policy offers great potential in providing timely fine grained security control on server and therefore enhance performance and security of e-commerce sites.

Keywords- Admission control, E-commerce, Overload control, Security, SSL Session.

I. INTRODUCTION

Security between network nodes over the Internet is traditionally provided using HTTPS i.e. SSL (Secure Socket Layer).It perform mutual authentication of both the sender and receiver of messages and ensure message confidentiality. This process involves certificates that are configured on both sides of the connection. Although providing these security increases remarkably the computation time needed to serve a connection, due to the use of cryptographic techniques. The disadvantage of the basic SSL approach is that security needs to be preconfigured on only predefined static information, namely the data and its location, that can be utilized when making renegotiation decisions. Adaptive SSL extends SSL beyond these limitations.

Adaptive security model was proposed in [1] which provide appropriate security mechanisms for SSL sessions at any particular moment in time as the environment changes. The adaptation controller for SSL called Adaptive SSL. concerned with adapting the choice of cryptographic algorithms applied to client-server interactions. [2] experimentally investigated the effects of security adaptation in various client-server scenarios .

In [3,4,5,6,7,8] various session based admission control (SBAC) mechanism for load management are discussed. Admission control is based on reducing the amount of work the server accepts when it is faced with overload. On most of the prior work, overload control is performed on per request basis, which may not be adequate for session-based applications, such as e-commerce applications. Session integrity, load management and security are the critical issue for the success of e-commerce site.

This paper proposed a frame work for enhancement of server and e-commerce application performance through load management using Adaptive SSL (ASSL) policy with session handling mechanism .This Adaptive SSL facilitates runtime adaptation of the SSL protocol by effectively intercepting requests between the client and SSL module and based on a set of conditions, may decide to renegotiate the session security i.e. selecting appropriate cryptographic algorithm at runtime. This self adaptive security will enhance performance of server as it reduces some of the overhead related to cryptographic algorithm used for secure connection.

The rest of the paper is organized as follows: Section 2 introduces the SSL protocol. Section 3 presents Overview of session based admission control techniques. Sections 4 introduces ASSL and Section 5 define proposed approach and Section 6 provide experimental environment detail and section 7 presents the conclusions of this paper.

II. SSL PROTOCOL

The SSL .protocol provides communications privacy over the Internet. It uses combination of public-key and private-key cryptography algorithm and digital certificates to achieve confidentiality, integrity and authentication. The SSL protocol increases the computation time necessary to serve a connection remarkably due to the use of cryptography to achieve their objectives. The study concludes that SSL connections is 7 times lower than when using normal connections as discuss in[7,8].

The four protocol layers of the SSL are Record Layer, ChangeCipherSpec Protocol, Alert Protocol, and Handshake Protocol ,they encapsulate all communication between the client machine and the server.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

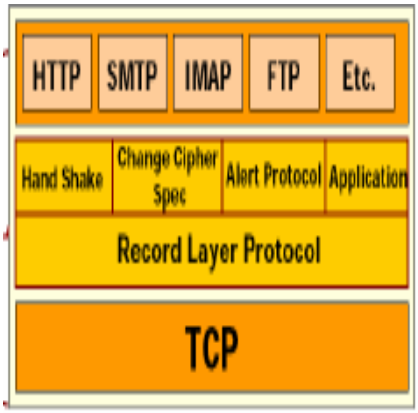


Fig 1 SSL Protocol

The SSL handshake allows the server to authenticate itself to the client using public-key techniques like RSA and then allows client and the server to cooperate in the creation of symmetric keys used for rapid encryption, decryption. Optionally the handshake also allows the client to authenticate itself to the server.

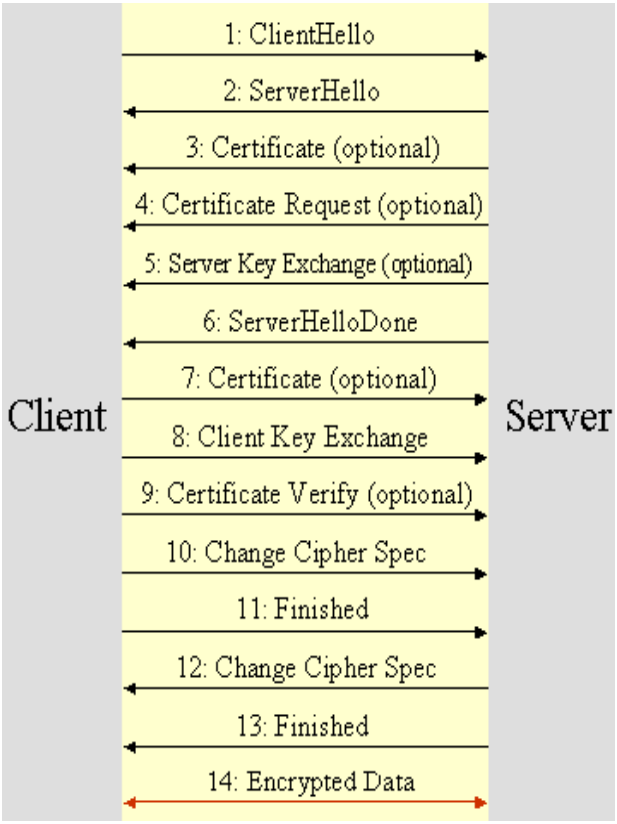


Fig 2 Handshake Protocol

Two different SSL handshake types can be distinguished: The full SSL handshake and the resumed SSL handshake. The full SSL handshake is negotiated when a client establishes a new SSL connection with the server and requires the complete negotiation of the SSL handshake, including parts that spend a lot of computation time to be accomplished. The SSL resumed handshake is negotiated when a client establishes a new HTTP connection with the server but using an existing SSL connection. As the SSL session ID is reused, part of the SSL handshake negotiation can be avoided, reducing considerably the computation time for performing a resumed SSL handshake. Notice that their is big difference between negotiate a full handshake respect to negotiate a resumed SSL handshake (175 vs. 2 ms) as discussed in[7,8].

SSL Connection

This is a logical client-server link, associated with the provision of a suitable type of service. In SSL terms, it must be a peer-to-peer connection with two network nodes. Every connection is associated with one session .

SSL Session

This is an association between a client and a server that defines a set of parameters such as algorithms used, session number etc. An SSL session is created by the Handshake Protocol that allows parameters to be shared among the connections made between the server and the client and sessions are used to avoid negotiation of new parameters for each connection. This means that a single session is shared among multiple SSL connections between the client and the server.

III. OVERVIEW OF SBAC TECHNIQUE

Following section give comparative study of various session based admission control mechanism for load management as discussed in [3,4,5,6,7,8]

CPU utilization based implementation presented in [3,4] is the simplest implementation of session based admission control but can break under certain rates and not work properly, reason is that the decision ,whether to admit or reject new sessions, is made at the boundaries of ac-intervals and this decision cannot be changed until the next ac-interval. However, in presence of a very high load, the number of accepted new sessions may be much greater than a server capacity, and it inevitably leads to aborted sessions and poor session completion characteristics

Hybrid admission control strategy covered in [5] which tunes itself to be more responsive or more stable on a basis of observed quality of service. It successfully combines most attractive features of both ac-responsive and ac-stable policies. It improves performance results for workloads with medium to long average session length.

Predictive admission control strategy also covered in [5] which estimates the number of new sessions a server can accept and still guarantee processing of all the future session requests. This adaptive strategy evaluates the observed

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

workload and makes its prediction for the load in the nearest future. It consistently shows the best performance results for different workloads and different traffic patterns. For workloads with short average session length, predictive strategy is the only strategy which provides both: highest server throughput in completed sessions and no (or, practically no) aborted sessions.

Session-based adaptive overload control mechanism based on SSL connections differentiation and admission control presented in [7,8] prioritizes resumed connections maximize the number of sessions completed and also limits dynamically the number of new SSL connections accepted depending on the available resources and the number of resumed SSL connections accepted, in order to avoid server overload.

IV. ADAPTIVE SSL

SSL protocol used to secure the communication channel between an application or web server and a client. The basic SSL secure transport layer connection, is established through a handshake mechanism where algorithms are selected based on those available to both the client and the server, for providing the three security properties confidentiality, authentication and data integrity. This process is commonly known as SSL negotiation and the resulting connection is called a session. Once established, the session can conduct a renegotiation,

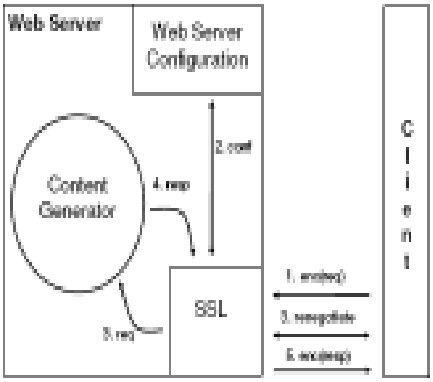


Fig 3 Basic SSL Model

(enc=encrypted, req=request, resp=response, conf=configuration)

Fig. 3 taken from [1,2] depicts a standard web server request-response processing cycle during an SSL secured session. The numbers in the figures indicate the event order and the labels the interaction type. Events with the same number indicate a decision point and only one of the events take place. It shows the client sending an encrypted request to the server in step 1. The request is initially passed to SSL which then queries the web server's configuration file (where the renegotiation decision rules are typically stored) and decides either to process the request or renegotiate the session. The disadvantage of the basic SSL approach is that security needs to be preconfigured only on predefined static information, namely the data and its location, that can be utilized when making renegotiation decisions.

Adaptive SSL extends SSL beyond these limitations. Previous work on Adaptive SSL (ASSL) covered in [1,2] facilitates runtime adaptation of the SSL protocol. ASSL aims to provide appropriate security mechanisms for SSL sessions at any particular moment in time as the environment changes. Attributing environmental factors could include the threat level, server load or transaction type. Client attributes such as processing power, bandwidth or type of client can also be considered. These factors are monitored by specialized third party applications that inform ASSL of affected clients and appropriate security measures. In [1,2] introduces a generic design for controlling the renegotiation within SSL and the resulting system called as 'Adaptive SSL'.

Design of Adaptive SSL

ASSL introduces a flexible approach to session management where renegotiation logic is decoupled from the main SSL implementation and web server configuration. Separating these concerns enables us to build a more powerful adaptive security model since the renegotiation rules can be determined and deployed independently and parallel to the web server and its components.

The ASSL module described in [1,2] effectively intercepts requests between the client and SSL module and based on a set of conditions, may decide to renegotiate the session security. The set of conditions are specified and altered at runtime by 'third parties' that is, other programs such as firewalls, system performance monitors, network monitors, server administrators, etc.. ASSL is used to create a self-adaptive security solution based on flexible use of renegotiation in SSL.

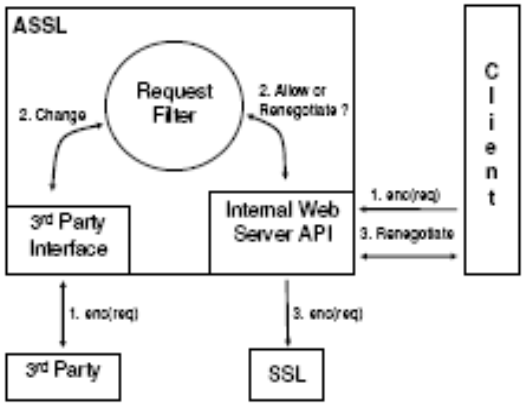


Fig 4 Adaptive SSL

Figure 4 taken from [1,2] explain how ASSL handles SSL requests. It effectively takes over the SSL negotiation and renegotiation logic by intercepting requests and evaluating them against the third party input. Decoupling the negotiation logic in this way from the main server logic, which is pre-configured in the "Web Server Configuration" allows to build

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

a more powerful and flexible adaptive security model.

Adaptive Policy

Adapting security based on current system performance is explored in [1,2] . In [1] it was shown that Adapting the security, i.e. changing the security algorithm at runtime has a server performance impact .It has investigated the different aspects of client load patterns.

In [2] a policy is defined to trade-off security and performance and shows how it performs in the scenario and finally experiment shows that the performance implications of adapting the security is highly dependant on the SSL session length and the requested file size. In [2] it was shown that performance cost incurred by the server through encryption and decryption is significantly influenced by the particular software implementation ,more costly algorithms should result in lower server throughput i.e 3DES is the most costly algorithm and is about 16 times more expensive than RC4 which is the least secure .

Table 1 Average number of bytes processed per second

Cryptographic Algorithm	***Throughput***
RC4	***112000***
AES 128	***36000***
AES 192	***30600***
AES 256	***26500***
DES	***19300***
3DES	***6900***

Based on this fact following adaptive policy is discussed in[2]

```
Pseudo code
#comment
load = getCpuLoad
PolicyTable(row, column)
curSec #current Security level
newSec #new Security level
#Move to X
Move to PolicyTable(curSec,curSec)
#Check if security should be increased
#Check all values to the left
IF load < PolicyTable(curSec, curSec - n)
RECORD newSec
#Check if security should be decreased
#Check all values to the right
IF load >= PolicyTable(CurSec, CurSec - n)
RECORD newSec
IF newSec THEN adapt
#end
```

This adaptation policy is very robust to client behavior. Firstly, security is reduced as soon as the server reaches its maximum load.. Secondly, security is only increased if the server could cope with the same load at a higher security level. When security is adapted only new clients start with the higher level of security and so it takes some time before the server is serving all clients at this new level .

Application Area

Applications for adaptive security have been proposed in a number of diverse areas. Examples include mobile ad-hoc networks where current network conditions play a role in choosing relevant security protocols at runtime . Media streaming research proposes the use of additional contextual information, in the form of an intelligent routing infrastructure, to allow for flexible security levels . There have also been proposals to provide adaptive system security through system event monitoring .

This Adaptive SSL research is preparatory work to create a self-adaptive security solution based on flexible use of renegotiation in SSL. Adaptive SSL policy is applied in different realistic scenario and establish strategies and optimizes the trade-off between security and performance (overhead as well as server performance) when renegotiating security level. The proposed work is discussed in following section.

V. *PROPOSED APPROACH*

The proposed work utilizes session handling mechanism with adaptive SSL policy for load management. This policy provides more flexible approach to session management where renegotiation logic is decoupled from the main SSL implementation and web server configuration. Runtime adaptation of cryptographic algorithm on online e-commerce site based on current number of active session help in increasing the number of completed session and enhancing server and application performance.

Phases of proposed framework

- Server Authentication
- Session Monitoring
- Application Of Adaptive SSL Policy
- Testing of Server and Application Performance

Phase 1: Server Authentication

Security between network nodes over the Internet is traditionally provided using Secure socket Layer .For communicating through SSL, server authentication is mandatory, the client tries to confirm the identity of the server based on the server's certificate. To accomplish this, on receiving a request from the client, the server sends its certificate to the client. This certificate contains information such as: server's public key, certificate's serial number, certificate's validity period, server's distinguished name,

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

issuer's distinguished name, and issuer's digital signature (a message signed using the issuer's private key). The client, on receiving this certificate, authenticates the server..

For performing the above authentication, the server must have a Public Certificate and Key store containing the key. A certificate can be self-signed when authentication over the internet is not really a concern, that is only data privacy and integrity are important .So self signed certificate is generate or certificate request is sent to CA for trusted certificate .So our work starts with generation of private and public key and self signed certificate and export it into server configuration file and hence configuring secure socket layer on server.

Phase 2: Session Monitoring

Session integrity is the critical metrics for e-commerce. So load management through session handling is performed which will enhance application performance. Firstly create a web application and deploy it on server. Then monitor the number of active session on that application and check against maximum limit as new session is created, if the maximum limit is reached , then do not allow the new session..

Phase 3: Application Of Adaptive SSL Policy

The proposal is to design a system for client server interaction which will monitor active sessions on the server and when the number of session reaches to alarming level then change highly secure cryptographic algorithm used for https communication to the algorithm which is less secure and consume less time in encryption and decryption process. This approach help in reduction of some load on the server due to less time consumption during cryptographic process. We are using security trade off between 3DES and RC4cryptographic algorithm. 3DES is more secured but gives least throughput and RC4 is the weak steam cipher but with good throughput capability.

Triple- Data Encryption Standard (3DES)

Triple-DES Using standard DES encryption, encrypts data three times and uses a different key for at least one of the three passes giving it a cumulative key size of 112-168 bits. The Triple-DES is considered much stronger than DES, however, it is rather slow compared to some new block ciphers hence give poor throughput as compared to other cryptographic algorithm.

RC4 Stream Cipher

RC4 is a cipher invented by Ron Rivest, co-inventor of the RSA Scheme.RC4 has two advantages over other popular encryption algorithms. First, RC4 is extremely fast. Second, RC4 can use a broad range of key lengths. For most ciphers, longer key length is better. However, RC4 was widely used primarily because its shortest optional key length is 40 bits. Unfortunately, RC4 is a dangerous cipher to use. If it is not implemented perfectly, its protection is minimal.

Phase 4: Testing of Server and Application Performance.

This phase will analyze our approach and show its importance in client /server paradigm. The main aim is to compare the throughput and response time of the application under different load pattern and in two different environment , which uses different cryptographic algorithm for security. This phase also provides visualization of the various performance metrics which can be utilized for further result analysis. In this phase there are two measures for evaluating performance (i) Throughput and (ii) Response Time

VI. CONCLUSION

This proposal of runtime adaptation of cryptographic algorithm using session handling mechanism will enhance throughput and response time of server . Session handling combines with Adaptive SSL is good alternative for load management. This work will provide self adaptive security control on server and optimizes trade off between security and performance.

ACKNOWLEDGMENT

The Success of this research work would have been uncertain without the help and guidance of a dedicated group of people in our institute MANIT Bhopal. We would like to express our true and sincere acknowledgements as the appreciation for their contributions, encouragement and support. The researchers also wish to express gratitude and warmest appreciation to people, who, in any way have contributed and inspired the researchers.

REFERENCES

[1] C. J. Lamprecht , Aad P. A. van Moorsel, “Adaptive SSL: Design, Implementation and Overhead Analysis”, prdc, pp.289-294, First International Conference on Self-Adaptive and Self-Organizing Systems (SASO '07), 2007.

[2] C.J. Lamprecht, Aad P. A. van Moorsel, "Runtime Security Adaptation Using Adaptive SSL," prdc, pp.305-312, 14th IEEE Pacific Rim International Symposium on Dependable Computing, 2008

[3] L. Cherkasova, P. Phaal “Session Based Admission Control: a Mechanism for Improving the Perfor-mance of an Overloaded Web Server.” HP Laboratories Report No. HPL-98-119, June, 1998.

[4] L. Cherkasova, P. Phaal “Session Based Admission Control: a Mechanism for Improving Performance of Commercial Web Sites.” prcd, Seventh International Workshop on Quality of Service, IEEE/IFIP event, London, 1999.

[5] L. Cherkasova, P. Phaal “Session Based Admission Control: a Mechanism for Peak Load Management of Commercial Web Sites.” IEEE J. Transactions on Computers, Vol. 51, No. 6, June 2002.

[6] M. Arlitt, “Characterizing Web User Sessions", ACM SIGMETRICS Performance Evaluation Review, Vol. 28, No. 2, pp. 50-56, September 2000.

[7] Jordi Guitart , David Carrera , Vincenc Beltran , Jordi Torres , Eduard Ayguade, "Session-Based Adaptive Overload Control for Secure Dynamic Web Applications" Proceedings of the 2005 International Conference on Parallel Processing, pp.341-349, 2005.
[8] Jordi Guitart , David Carrera , Vicenç Beltran , Jordi Torres , Eduard Ayguade, "Designing an overload control strategy for secure e-commerce applications," Computer Networks: The International Journal of Computer and Telecommunications Networking, v.51 n.15, p.4492-4510, October, 2007
[9] P. Lin, "So You Want High Performance" (Tomcat Performance), September 2003, online available URL: http://jakarta.apache.org/tomcat/articles/performance.pdf.
[10] Sun Microsystems, "Java Secure Socket Extension (JSSE)." online available URL: http://java.sun.com/products/jsse/.
[11] Jakarta Project. Apache Software Foundation, Jakarta Tomcat Servlet Container. URL: http://jakarta.apache.org/tomcat.
[12] A.O. Freier, P. Karlton, C. Kocher, "The SSL Protocol Version3.0" November 1996. available online URL: http://wp.netscape.com/eng/ssl3/ssl-toc.htm

R K Pateriya M.Tech & B.E. in Computer Science & Engg. and working as Associate Professor in Information Technology Department of MANIT Bhopal . Total 17 Years Teaching Experience (PG & UG). Guided twenty M.Tech Thesis .

Dr. J. L . Rana Professor & Head of Computer Science & Engg deptt. in MANIT Bhopal .He has received his PhD from IIT Mumbai & M.S. from USA (Huwaii) .He has Guided Six PhD.

Dr. S. C. Shrivastava Professor & Head of Electronics Engg. department of MANIT Bhopal. He has Guided three PhD , 36 M.Tech and presented nine papers in international & twenty papers in national conference in India.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

An Enhancement on Mobile TCP Socket

S.Saravanan,
Research Scholar
Sathyabama University
Chennai, India

Dr.T.Ravi
Professor & Head, Dept. of CSE
KCG College of Technology,
Chennai, India

Abstract – ***A TCP session uses IP addresses (+ IP port) of both end points as identifiers. Therefore when a mobile handover to a new AP that belong to a different subnet/domain, the IP address will changes and ongoing TCP connections are reset. Several approaches have been proposed to solve this problem, and one of which was to modified the TCP/IP stack to update the changes of the IP address for the ongoing connections [5] [6]. However, these proposals causes unnecessary processing when TCP is used in applications which have already employed some kinds of security measures, such as SIP. This paper proposes the Mobi Socket, which specifically supports TCP mobility for intrinsic secure applications without unnecessary overhead.***

1. INTRODUCTION

TCP/IP was developed when all network nodes were stationary, and connection to a network is through cable, therefore it is unthinkable that a node will move to another subnet while ccessing to the Internet, and the IP address of an end host is assumed to stay unchanged while a computer is running. As a result, IP address (together with IP port) is used as identifiers for TCP session, and the TCP layer at the end host maintains TCP control blocks (TCBs), which hold the IP addresses and IP ports of both ends for each TCP connection to fmd the right socket for each datagram it receives from the IP layer.

But with the introduction of wireless access technologies such as Wi-Fi, it is possible for a mobile node to handover to a new AP that belong to a different subnet/domain while actively connecting to the Internet. This causes an IP address change, and for current implementation of TCP/IP stack, all ongoing TCP connections are reset. This will cause problem for long-live TCP sessions.

There are two general approaches to solve the problem of changing host IP address for TCP session. The first one uses the split-connection approach, which introduces a fix middle agent between the mobile host (MH) and correspondent host (CH) [4]. The connection between CH and MH is broken into two parts, the fixed part between the CH and the agent remains unchanged regardless of the position of the MH, and the TCP connection between the agent and the MH will be re-established whenever the MH handovers to a new address. In this sense only the TCP at the MH is affected, while at the CH the TCP session is not disturbed. The problems of these approaches are non-transparent end-to-end operation of TCP session, as well as the requirement of new infrastructure entities (the middle agent) and triangle overhead.

The other approach modify the TCP stack so that when the mobile host changes the connection to the internet, the TCP stacks at both ends preserver the TCP connection and update the TCBs with the new IP address at both ends accordingly.

In [5], when the MH changes its location, the proposal in [5] introduced new states to the TCP specification. When the address of MH changes, MH and CN will exchange information and update the new IP address accordingly. Both sides will prepare in advance a share-secret, and use this sharesecret to authenticate each other during the update process.

The proposal in [6] employs a similar concept, but instead of changing the TCP stack, it uses kernel extensions and a userlevel redirect daemon process (this was the design of the prototype in BSD). The daemon process will monitor the wireless network interface for changes of IP address, and if one is detected, the daemon at the MH will inform the counterpart at the CH to update the new IP address together. To secure the update process from malicious acts, MH and CH also need a share-secret in advance.

The problem with [5] is that both sides has to perform additional works to exchange a share-secret in advance, regardless of whether the MH will actually performs the handover to a new Access Point (AP) or not, or whether the TCP session lives long enough to experience a handover. The proposal in [6] relieves this matter by initiating the preparation process only if the TCP connection exists longer than a threshold. However, if the MH does not perform a handover, then all of the preparations for the long-live connections are wasted.

One more problems with [5] and [6] is that processing the share-secret for authentication will requires a lot of processing, which in tum consumes battery power at the MH. If many TCP connections are used (such as if the user constantly browsing the Internet) then battery life will be shortened considerably. Moreover, both [5] and [6] are not applicable in the case where both ends perform handover simultaneously.

In the next parts of this paper, we propose a new type of socket called the TCP MobiSocket, that remains connected even if the concerned IP address changes. It works like normal TCP socket, but does not get reset when the IP address at either end changes, and with an additional updateTCB() member function to update the TCBs with the new address. All of the security issues that are required to secure the update of the new address will be handled by the calling applications. This new socket is dedicated to support

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

mobile TCP session for intrinsically secure application, without all the above mentions problems of[5] and [6].

2. DESIGN OF THE MOBISOCKET

Logically, there are two phases when mobile device handover. First, the Network interface/card disconnects from the old AP. Then it connects to the new AP. In traditional TCP stack, the network stack at the MH will close all TCP connections in cleaning-up activities, as well as reset the TCBs during these phases.

On the other hand, all of the ongoing MobiSocket will remain in ESTABLISHED state, when the IP address changes, waiting to be updated by the application.

We design a new socket that allows the application to update the change of PoA at both end hosts. The socket is designed based on the following assumptions/requirements:

- There are cases when the TCP connection needs explicit handling before communicating using the new IP address (Re-establishment/update of Security Association for VPN, sending the PATH message of RSVP for QoS, etc...)
- The application takes care of security activities regarding the update of the new address. The reasons for this are (l) if the connection needs to be secure, the applications have already shared some kind of security, and (2) if the connection is not important to the extend that it requires a shared security association between both end host, then it might not important enough to be hacked by others.
- Compatibility with applications using legacy TCP/IP stack is desired to promote deployment. It means that in the case the other end does not support the features of mobile socket, connection will work according to that of legacy TCP specification.
- Being able to provide handover of TCP session between different network interfaces of the same mobile device. The requirement is that not only IP address but change of TCP port also must be supported, because the same port of the other network interface might be in use by a different application at the time of the request for handover.

The application which uses the MobiSocket will call the MobiSocket's updateTCB() member function to update the TCB with the new destination address. To satisfy the above requirements, the mobile socket will provide the following APIs to the applications:

- acvMobi(socket_id)
  socket id: the handler of the socket
  - The application will call this function to explicitly activate the mobile feature of the socket
  - If this function is not called, then the MobiSocket will work like normal TCP socket
  - When this function is called, the TCP connection will not be abolished if the concerned wireless interface changes to a new IP address
- updateTCB(socket_id, direction, newIPaddress, newPort)

socket id: handler of the socket
direction: update the source or the destination address
newIPaddress, newPort: the new IP address and new port to update to TCB/PCB (TCP Control Block/Protocol Control Block). If the port is 0 then keep the existing port value

  - The application will call this function to update the TCBIPCB (TCP Control Block/Protocol Control Block) with the new source/destination address and port
  - The MobiSocket will start a new congestion control algorithm called the mobile congestion control
- copyTCB(new_socket_id, old_socket_id)

old socket id: handler of the old socket
new- socket- id : handler of the new socket

  - The application will call this function to update the TCBIPCB (TCP Control Block/Protocol Control Block) of the newly created socket with the information of the old socket. This is used when the application want to handover from old interface to new interface.
  - This function will copy all information of the old socket (include current states, CWND, AWND, RTO etc..., except the source IP address and source Port) to that of the new socket, and then delete the old socket without sending FIN to the other end (i.e., application at the CH).

Apart from the above two new APIs/member functions, the MobiSocket also introduces two new message, the AddChange and AddConfirm.

The AddChange contains (I) A shared token between Mobile Host (MH) and Correspondent Host (CH), (2) the old IP address and the new IP address encrypted by the private key of the MH, (3) the new port address and (4) The old IP address of the MH in plain-text.

The AddCorifirm contains (l) the shared token between MH and CH, (2) the new IP address encrypted by the private key of the CH.

If the two messages above are implemented as TCP header options, then these header options must be sent to the applications, but currently there is no mechanism to perform such action. Therefore, it might be better to send this as OOB (out-of-band) data using the TCP Urgent Pointer .

3. WORKING PROCEDURE OF THE MOBISOCKET

Let's consider the use of MobiSocket for a SIP application. Suppose that a TCP connection is established between MH and CH (the thick, solid line), which have established a SIP session through the SIP server. The MobiSocket will work as follows (see figure 1):

- First the application creates the TCP socket for the SIP session, and calls the acvMobi () to activate the mobile feature for the socket

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

- In step <1>, the MH moves from Subnet 1 to Subnet 2, and in the process its address change from IPaddress 1 to IPaddress2
- In step <2>, the SIP application at the will call the updateTCB () function to replace IPaddressl with IPaddress2 at the TCB table. Then it issues a SIP INFO message to ask CH update the new IP address of the MH.
- Upon receiving this INFO message in step <3>, the SIP application at the CH will authenticate the message using SIP security associations, and all the updateTCB () function to replace IPaddressl with IPaddress2 at the TCB table.
- Then in step <4> the SIP application at the CH will send back the INFO message back the the MH to confirm the change of address. Note that both INFO messages may contain other parameters of the concerned TCP session, such as new window size, MSS etc ...
- CH and MH will start sending data using the TCP connection when they receives the INFO message from the other end, and they will start receiving data after they send the INFO message to the other end.

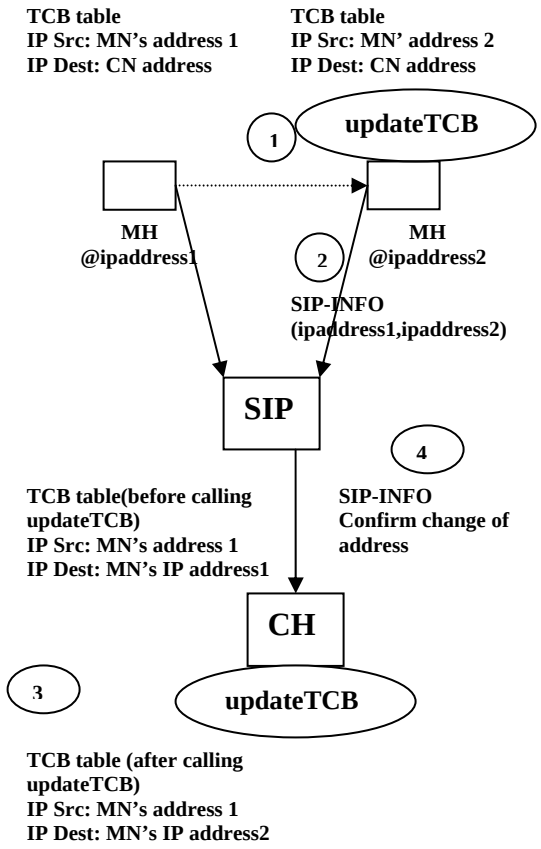


Figure 1: Working procedure of the mobisocket .

4. DISCUSSION AND OPTIMISATION

The merits of the MobiSocket are:

- Inherit intrinsic security feature of SIP
  - Less processing overhead for security issues (conserve power)
- Depending on the security requirements, the application can decide whether to allow the handover of TCP connection
  - More suitable for application with strict security requirements
- Still work when both ends handover simultaneously
  - Reach-ability through SIP Registration

functionality

We can further optimize the operation of the MobiSocket as follows:

When the MN receives the INFO message from CN, both ends might already time out (due to handover, NOT due to congestion), so even if the TCB is updated, no data exchange is possible until the time out is over (can be very long). We can provide a new function to reset the timer after the updateTCBO function, which is the reset'I'imensocket id). This function will reset the TCP socket to the state as if it has just received a datalACK packet from the other machine.

Furthermore, if SIP proxy is used, then normally the MN has to finish re-Registration with the SIP proxy first before it can send SIP INFO message to the other end. This creates further delay for the TCP session. To solve this, we note that the MN and CN can share public key with each other during the initial INVITE process, therefore after the MN handover to a new IP address, it can use the public key of the CN to send the SIP INFO message to the CN right away. However, this solution cannot be used if both ends handover simultaneously (therefore they do not know the IP address of each other), in this case they must contact through SIP proxy server (after the re-Registration process)

5. CONCLUSION

In this paper we propose the MobiSocket to support TCP mobility for secure application such as SIP. This socket causes no overhead if handover does not take place like previous proposal, and moreover it still works when both side handover simultaneously.

In this socket, there is no need for per-TCP connection authentication, because the authentication is left to application. Depending on the real situation, the application can also control whether to keep the TCP session or not, which is more appropriate for application which is applied with other application level constrains such as security and QoS policy .

In the future, we would like to carry out the implementation of the MobiSocket to confirm the design of the system, as well as to measure the delay and throughput parameter when the reset'Iimen) function is (1) called and (2) not called, and compare the results with that of [5] and [6].

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

We also would like to measure the delay in the case of SIP application, when we send the INFO message before and after re-Registration, as well as when two end hosts handoff together.

We also plan to update the proposal in [1] with this new MobiSocket.

REFERENCES

[I] Vu Truong Thanh, Yoshiyori Urano, "Agent based LLMA handover scheme for SIP communication - The case for UDP traffic" , The II th International Conference on Advanced Communication Technology (rCACT), Feb. 2009

[2] C. Perkins, "IP Mobility Support for IPv4", Request for Comments: 3344, IETF, August 2002

[3] Huei-Wen Ferng et. ai, "A SIP-Based Mobility Management Architecture Supporting TCP with Handover Optimization", Proc. Of Vehicular Technology Conference, pp. 1224-1228, Apr. 2007

[4] Milind Buddhikot et. ai, "MobileNAT: a new technique for mobility across heterogeneous address spaces", Proc. the 1st ACM international workshop on Wireless mobile applications and services on WLAN hotspot, pp. 75-84, Sept., 2003

[5] FUNATO D., "TCP-R: TCP mobility support for continuous operation", Proc. IEEE International Conference on Network Protocols, pp.229 -236 ,Oct. 1997

[6] Vassilis Prevelakis and Sotiris Ioannidis, "Preserving TCP Connections Across Host Address Changes", Lecture Notes in Computer Science, Springer Berlin / Heidelberg, pp. 299-310 Oct., 2006

[7] Rosenberg, et. aI., "Session Initiation Protocof', Request for Comments: 3261, IETF June 2002

[8] D. Yon et. ai, "TCP-Based Media Transport in the Session Description Protocol (SDP)", Request for Comments: 4145, IETF, September 2005

AUTHORS PROFILE

S. Saravanan B.E., M.E., (Ph.D) working as an Assistant Professor at Jeppiaar Engineering College, Chennai and he has more than 9 years of teaching experience. His areas of specializations are Computer Networks, Network security and TCP/IP.

Dr. T. Ravi,B.E,M.E,Ph.D is a Professor & Head of the Department of CSE at KCG college of Technology, Chennai. He has more than 18 years of teaching experience in various engineering institutions .He has published more than 20 papers in International Conferences & Journals.

Technology in Education

Modern Computer Graphics Technologies Used at Educational Programs and Some Graphical output screens

[1]N. Suresh Kumar, [2]D.V. Rama Koti Reddy, [3]S. Amarnadh, [4]K. Srikanth, [5]Ch. Heyma Raju,
[6] R. Ajay Suresh Babu, [7]K. Naga Soujanya
[1, 3, 4, 5] GIT, GITAM University, Visakhapatnam
[2] College of Engineering, Andhra University, Visakhapatnam
[6]Raghu Engineering College, Visakhapatnam
[7] GIS, GITAM University, Visakhapatnam
[1]nskgitam2009@gmail.com

Abstract In This paper a new technique is implemented to teach microprocessor and to clarify the doubts in the subject microprocessor. Although a lecturer has many aids to explain the topic in the class room, but a graphical environment is more power full environment in the present education scenario, which can improve the student level of understanding. The graphical interface develops the concepts of the student graphics concepts and also a student can easily grasp any level of task. The lecturer shows a visual object or animated show to the student to explain the particular topic in the class room. This work includes the framework of graphics programming; students can concentrate on the technical subject. Thus they acquire a method to construct computer graphics programs in many ways and gain knowledge in the concerned technical paper. The project have used for six years, and convinced of the positive effect.

I Introduction

E-learning substantially improves and expands the learning opportunities for students [3]. The modern computer information technologies, which are widely used both at educational programs for conducting of effective lecture, conducted scientific researches, and forming of practical and laboratory works with the students of technical and computer-based special[2]. Every teacher comes to understand that successful imparting of information and skills lies in the ability to incorporate a variety of technologies that, directly or indirectly, help communication between student and teacher [4]. Advances in learning objects and other technologies "that will optimize interoperability with other institutions and organizations in areas such as the creation of learning objects databases, information databases such as libraries, administrative systems and learner support strategies as well as the facilitation of interactions among learners and teachers", will continue to expand the scope of possibilities with which educational institutions will have to grapple [5].

II Experimental Idea

The project was to develop software delivered diagnostic and remedial system for Diploma and Graduate students taking subjects in microprocessors and embedded systems. The system is comprised of a module to diagnose a student's background microprocessor knowledge and modules to which the student is directed to remedy any deficiencies that are found. The remedial modules focus on common areas of weakness in microprocessor architectures and interfacings and revise the basic peripherals that are needed to interface. For example, in memory interfacing the number of address lines and microprocessor type are incorporated with the graphic tools such as colors, sounds, messages and animations as shown in figure 1, 4, 5. When a student select any part on the selected architecture it will link to the particular file which give the details of the unit in the architecture as shown in figure 2, 3. The Interfacing devices like 8255 and it modes can be set in control register (CR) by an interactive message box as shown in the figure6. The software also demonstrates some of the bio-Medical applications and some interfacing designs such as traffic light controllers, stepper motor controllers, seven segment interfacing etc. This software allows the user to develop any interfacing circuit interactively as shown in figure 4. For maximum accessibility and ease of implementation, the software was designed to run in a Windows environment using graphics application in C language. This allows the software to be used off-campus, and makes the system equally accessible to students who are studying by distance education, open learning or who are located on an overseas campus [1].

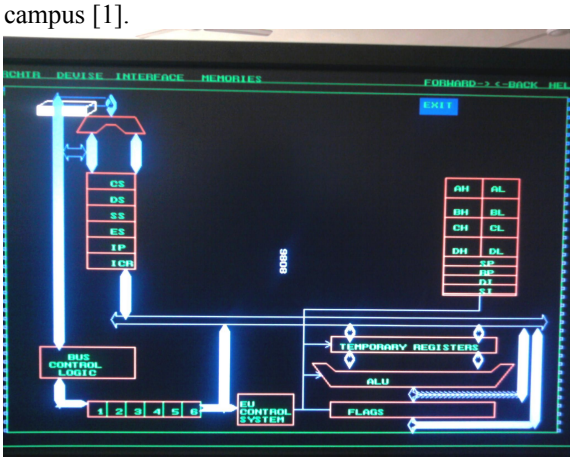


Figure 1 8086 Architecture

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

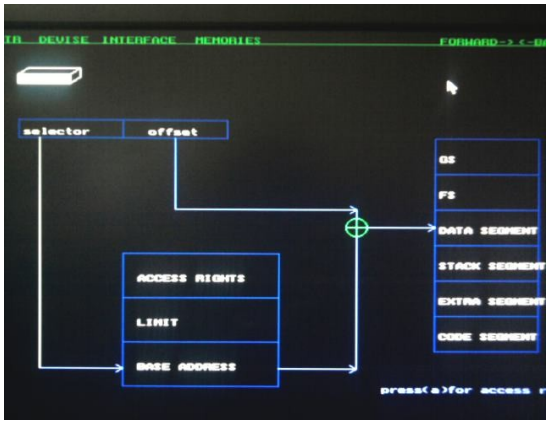


Figure 2 Segment address

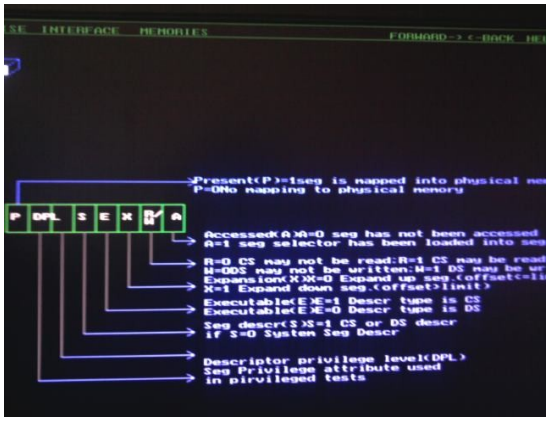


Figure 3Access rights and Permissions

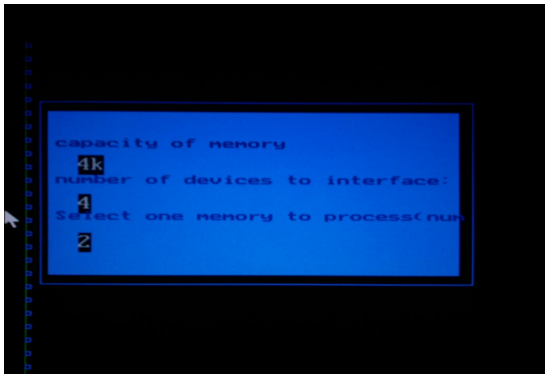


Figure 4 Dialog box permits to set the memory settings by user

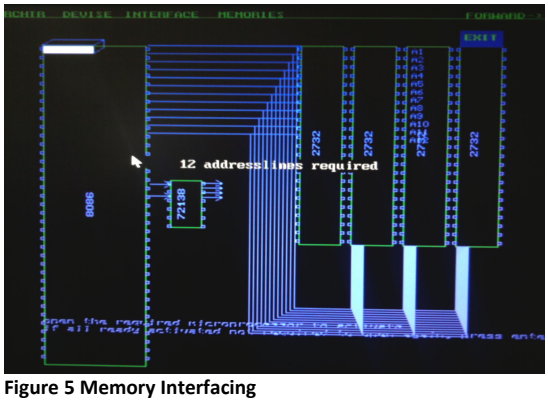


Figure 5 Memory Interfacing

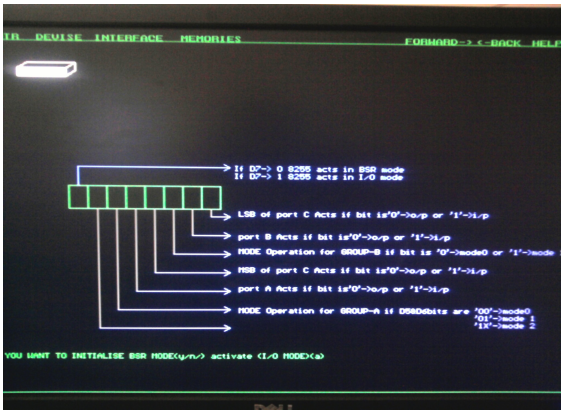


Figure 6 8255 CR msg box which allow the user to set the mode operations

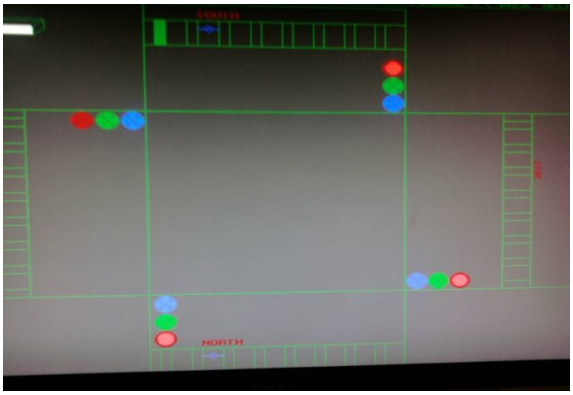


Figure 7 Traffic Light Demonstration

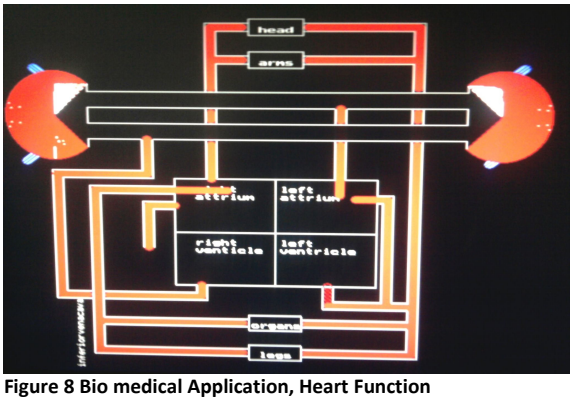


Figure 8 Bio medical Application, Heart Function

Deficiencies in background knowledge and skills need to be diagnosed early before they become barriers to students' further learning in microprocessor subjects. it is not possible to offer individual tuition. The problem is exacerbated with off campus students taking Distance Education or Open Learning subjects. The software also help teaching staff relieved from from his duties for some time where he can attend other problems and enabling them to concentrate on presenting the primary subject material. The anticipated outcome will be improved student attitudes, and lower dropout and failure rates in a subject area which is difficult for many. Wade and Power [6] suggest the following "General Requirements for WWW Based Instructional Design:"

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

1) The presentation of material should support a range of sensory experiences incorporating interactivity and multimedia elements.
2) Students should be provided with the opportunity to experiment with the knowledge they have learned.
3) Testing and checkpoints are important from the point of view of repetition and student retention.
4) Educational software should motivate the student.
5) The learning environment should support the cognitive structures of the student;
6) Facilities for synchronous communication and collaboration should be supported where possible.
7) A well-designed interface will enable the student to interact with the material without the complex intermediaries and will aid in the understanding of the knowledge domain and structure.
8) The development of a Tele-Educational course requires the support and cooperation of faculty and administration.
9) WWW-based educational courses must be integrated into a well understood and explicitly specified curriculum this includes clear objectives, content description, methods of teaching, student learning, and student assessment and course evaluation.

In the present paper the work includes two stages
Stage 1:

- Develop soft ware to diagnose the student
- Develop assignments
- Seek the student feed back
- Ask the student to create his own assignments
- Conduct quiz and descriptive test

Stage 2:

- Make interactive session to assess the student
- Seek the student feed back
- Propose alterations and development in the soft ware if any

Preferably, these procedures are conducted early enough at the end of each unit or chapter in the semester to allow weak students to take remedial action before being challenged. Conducting these stages most of the student requirements for WWW Based Instructional Design are satisfied which are proposed by wade and power [6].

III Features

- The structure of the software especially the presentation, animation, colors, fonts and sound affects made interest in students to diagnose him.
- The software structure is very interactive and attractive.
- By further development of the modules in the soft ware, it can use for research and real time application problems.

IV Requirement

Our software package is developed in graphics using C – language in Turbo C++ editor. The software is developed on Microsoft window – 95 based system. The main aim is, in poor countries low economic colleges and lagging in new technologies can also use this software. The systems require a graphic card and program is developed for screen resolution 640X480.

V Conclusion

On the basis of comparative analysis the software packages is marked as versatile direction in conducting lectures, lab work and diagnose weak student performance. E-learning substantially improves and expands the learning opportunities for students. **The modern computer information technologies, which are widely used both at educational programs for conducting of effective lecture, conducted scientific researches, and forming of practical and laboratory works**
Preferably, these procedures are conducted early enough at the end of each unit or chapter in the semester to allow weak students to take remedial action before being challenged. Conducting these stages most of the student requirements for WWW Based Instructional Design are satisfied

References

[1] Max King, Ian Kirkwood, Clive McCann, Department of Econometrics, Monash University," The soft(ware) approach to learning problems in quantitative methods" 0-7803-3173-7/96/$5.00 @ 1996 EEE
[2] Vasyl Zayats, Vasyl Kogut, "Role of Information Technologies in Progress of Science and Education" MEMSTECH'2009, 22-24 April, 2009, Polyana-Svalyava (Zakarpattya), UKRAINE.
[3] Robert S. Friedman and Fadi P. Deek, "Innovation and Education in the Digital Age: Reconciling the Roles of Pedagogy, Technology, and the Business of Learning" IEEE Transactions on engineering management, VOL. 50, NO. 4, NOVEMBER 2003,pg 403
[4] R. Friedman and F. P. Deek, "Problem-based learning and problemsolving tools: synthesis and direction for distributed education environments," *J. Interact. Learn. Res.*, to be published
[5] F. P. Deek, M. Deek, and R. Friedman, "The virtual classroom experience: Viewpoints from computing and humanities," *J. Interact. Learn. Environ.*, vol. 7, no. 2/3, pp. 113–136, 1999.
[6] V. Wade and C. Power, "Evaluating the design and delivery of WWW based educational environments and courseware," in *Proc. ACM-ITiCSE*, Dublin, Ireland, 1998, pp. 243–248.

Impact of language morphologies on Search Engines Performance for Hindi and English language

Dr. S.K Dwivedi
Reader and Head, Computer Science Dept.
BBAU
Lucknow, India
skd200@yahoo.com

Rajesh Kr. Goutam
Research Scholar, Computer Science Dept.
BBAU
Lucknow, India
Rajeshgoutam82@gmail.com

Parul Rastogi
Research Scholar, Computer Science Dept.
BBAU
Lucknow, India
parul_rastogi@yahoo.com

***Abstract* - Search Engines are the basic tools of Information Retrieval on the web. The performance of the search engines are affected by various morphological factors of the language. The paper covers the comprehensive analysis and also the comparison of the impact of morphological factors and other language structure related factors (like sense ambiguity, synonyms) on the performance of Hindi and English language information retrieval on the web.**

Keywords: Hindi language morphology; English language morphology; web searching; morphological structure; precision; Hindi language Search Engines.

I. INTRODUCTION

Hindi language is the national language of India, with roughly 300 million native speakers. Another 100 million or more use Hindi as a second language. It is the language of dozens of major newspapers, magazines, radio and television stations, and of other media.

In Hindi language, nouns inflect for number and case. To capture their morphological variations, they can be categorized into various paradigms based on their vowel ending, gender, number and case information.

Hindi language Adjectives may be inflected or uninflected, e.g., `चमकीला' {chamkiila} (shiny), 'अच्छा' {acchaa} (nice), `लंबा' {lambaa} (long) inflect based on the number and case values of their head nouns while `सुंदर' {sundar} (beautiful), `भारी' {bhaarii} (heavy) etc. do not inflect. Hindi Verbs inflect for the following grammatical properties Gender, Number, Person, Tense, Aspect and Modality (GNPTAM).

The morphemes attached to a verb along with their corresponding analyses help identify values for GNPTAM features for a given verb form [1]. Hindi language has core distinctive structure which affects the results of the search queries on the web.

English is also the popular language in India and across the globe. Though, the morphological structure of English language is not complicated as Hindi but it also follows some of the similar morphological constraints.

Morphology is the study of the structure of words. The structure of words can also be studied to show how the meaning of a given morpheme, or its relation to the rest of the word, varies from one complex word to another. Consider how sun works in the following words: sunbeam, sunburn, sundial, sunflower, sunglasses, sunlight, sunrise, sun-spot (scientific sense), sun-spot (tourist sense), suntan. Inflection does not really yield "new" words, but alters the form of existing ones for specific reasons of grammar. Derivation, on the other hand, does lead to the creation of new words.

Sometimes a word is changed in its form to show the internal grammar of a sentence. Examples would be plural forms of nouns (dog + s → dog-s) or past (imperfect) tenses of regular verbs (want + ed → want-

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

ed). The study of such changes is inflectional morphology (because the words in question are inflected - altered by adding a suffix).

Other compound or complex words are made by adding together elements without reference to the internal grammar of a sentence. For example, the verb infect suggests a new verb disinfect (=to undo the action of infecting). New words are often formed by noun + -ize, noun + ism, or verb + able (scandalize, Stalinism, disposable). The study of such words, "derived" from existing words or morphemes is derivational morphology. Like Hindi language the morphological structure of the English language also has an impact on the performance of the search engines.

II. SEARCH ENGINES SUPPORTING HINDI LANGUAGE

With India emerging as a highly competitive Internet market, more and more people are turning towards the Internet. The number of internet users in India is increasing day by day. Now it is high time that India should have its own search engine leaning towards Indian language contents. Few of the search engines are already there in the market like- Google, Raftaar and Guruji. Google [2] is capable of searching information in many Indian languages successfully. Google is far better then other Hindi supporting search engines. Raftaar[3] is developed by Indicus Labs. It is said completely Hindi language search engine. Funded by Sequoia Capital, developed by students of IIT, Delhi Guruji [4] is also one of the popular search engines for Indian languages. Guruji Search can be described in one word - fast. The search results are thrown up pretty darn quickly with all the search results pointing towards note worthy Indian sites.

III. FACTORS AFFECTING PERFORMANCE OF SEARCH ENGINES

The information retrieval on the web in any language faces numerous challenges. Besides all the technical factors the morphological structure of the language is one of them which also affect the performance of the information retrieval on the web.

A. Morphological factors

Morphological factors are related to the morphology of any language. As discussed in previous sections the morphological structure of Hindi and English language is somewhat similar to an extent. The Hindi as well as English language faces the similar challenges in Natural Language Processing tasks, which also includes Information Retrieval on the web. Following are the challenges which both the language faces while information retrieval on the web.

1) Root word of the keywords: Every language use some markers like (English language use s, es, ing and ें, ों, यॉ in Hindi language) are used with a root word and new words are constructed. These new words are called morphological variants of the stem. For ex. नदियाँ {Nadiyaan} (rivers), नदियों {Nadiyon} (rivers) are morphological variant of root word नदी{Nadi} (river). While searching on the web users give numerous queries which consist of words which are not used in their root form.

Similarly in English language also some end markers are used to constuct new words as already discussed in this paper. For ex. Exam and Examination, Percent and Percentage, River and Rivers etc.

The process of stemming of word which converts the word into its canonical form/ root word is obligatory. It is observed from th results that sometimes it improves the performance of the search engines.

We had taken a sample set of 50 TREC queries for English language and 50 Hindi language queries collected from the various users to test the affect of the root word on the performance of the search engines.

Following Table 1 and 2 are the set of randomly selected queries from this set. We compared the performance of the English and Hindi language search engines in the light of the root word. For Hindi language we had tested our queries on Google, Raftaar and Guruji. For English language, we had tested queries on Google, Altavista and Yahoo.

Query# 1 समाज में स्त्रियों की स्थिति {Samaaj mein striyon ki sthiti} (women's situation in community)
Morphological variants of Query#1 are as follows -
#1.1 समाज में स्त्रियों की स्थिति, #1.2 समाज मे स्त्री की स्थिति

Query#2 धार्मिक विवादों का अन्त {Dhaarmik vivadon kaa ant} (end of religious disputes)
Morphological variants of Query#2 are as follows -
#2.1 धार्मिक विवादों का अन्त, #2.2 धार्मिक विवाद का अन्त

Query#3 धर्म पुराणों {Dharm puraadon} (Religious Pandora)
Morphological variants of Query#3 are as follows -

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

#3.1 धर्म पुराणों, #3.2 धर्म पुराण

Query#4 भारत के राज्यों {Bhaarat ke Rajyon} (States of India)

Morphological variants of Query#4 are as follows -

#4.1 भारत के राज्यों, #4.2 भारत के राज्य

From the results it is evident that only Google indexes the documents keyword in their root form. Raftaar and Guruji do not index in that form that is the reason number of document retrieved is also less in comparison to Google. The overall comparison of the results from the three search engines show that in general the quantity of result retrieved increased when the keywords are used in their root form.

In the case of search engines the quality of results is more important than the quantity of the results. Table 1 shows the comparison of the precision value of the three search engines. The precision value is calculated by taking top ten results of the search engines. On closely observing the result, we can say that the precision value in the case of Google is high in almost all the queries. We observed that relevancy of the results is also high in Google in comparison to other two search engines, Raftaar and Guruji which denotes that not only quantity but the quality of the result is also affected by the morphological variations of the keywords.

TABLE I. PRECISION VALUES OF GOOGLE, RAFTAAR AND GURUJI AGAINST HINDI LANGUAGE SAMPLE SET QUERIES

Query#	Results for (Google)		Results for (Raftaar)		Results for (Guruji)	
	Doc. Retrieved	Pre.	Doc. Retrieved	Pre.	Doc. Retrieved	Pre.
1.1	1,87,000	0.90	1,322	0.50	1,87,000	0.80
1.2	34,200	1.0	3,230	1.0	34,200	0.80
2.1	1100	0.20	647	0.20	68	0
2.2	1000	0.20	1379	0.50	844	0
3.1	19,100	1	2,903	0.5	7760	0.4
3.2	24,100	1	2,609	0.8	21,075	0.5
4.1	6,92,000	1	61,691	0	1,55,543	0.3
4.2	6,06,000	1	2,24,911	0	6,16,188	0.3

Similarly we had taken five sample queries for English language also to test the affect of Root Word on the performance of the Search Engines.

Query#1.1 Civil service exam 1.2 Civil service examination

Query#2.1 Water wastage in India 2.2 Water waste in India

Query#3.1 Funding and grants institution 3.2 Funds and grants institution

Query#4.1 Mercury levels in birds 4.2 Mercury level in birds

TABLE II. PRECISION VALUES OF GOOGLE, ALTAVISTA AND YAHOO AGAINST ENGLISH LANGUAGE SAMPLE SET QUERIES

Q#	Results for (Google)		Results for (Altavista)		Results for (Yahoo)	
	Doc. Retrieved	Precision	Doc. Retrieved	Precision	Doc. Retrieved	Precision
1.1	1,340,000	0.71	31,400,000	0.57	30,500,000	0.64
1.2	5,590,000	0.66	31,400,000	0.57	1,600,000	0.68
2.1	43000	0.68	698,000	0.47	1,660,000	0.44
2.2	2,390,000	0.52	29,400,000	0.64	3,260,000	0.68
3.1	62,300,000	0.26	25,800,000	0.67	25,600,000	0.36
3.2	78,800,000	0.54	21,800,000	0.46	21,600,000	0.45
4.1	4,860,000	0.56	1,790,000	0.53	1,640,000	0.67
4.2	9,180,000	0.59	3,480,000	0.45	2,850,000	0.46

It is observed that all the three search engines returned almost same top results links with major differences in coverage area. Only AltaVista sometimes shows the same coverage area in morphological variant words. However Yahoo and Google did not return similar results in morphologically variant words. The relevance level of four morphological variant queries is calculated for the three search engines. It is found that Google outperforms the other two.

2) Part of Speech (POS) Disambiguation: Part of Speech Disambiguation is also one of the problems while searching in Hindi language on the web. In Hindi language there are words that can behave as noun in some cases and the same word can behave as verb in some other case. It is desirable from the search engines to disambiguate the POS of the keyword in particular results. Ambiguous POS of a word will have a drastic affect on the results. For ex. Query 'गया शहर' {gaya shahar} {Gaya city) in which 'गया' is an ambiguous POS it is used as verb in the sense of 'gone' and also as noun in the form of city name. All the three search engines (Google, Raftaar and Guruji) are returning maximum results in the sense of 'gone'. It is desired from the search engines to disambiguate the POS of the keywords which are ambiguous in nature.

The similar problem was noticed in English language information retrieval on the web. For English

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

language we had tested the following queries on the three search engines
Query#1Code of conduct (noun)
Query#2 Conduct (verb) inaugural ceremony

It is observed that only Google is able to differentiate between noun and verb 'conduct', whereas Altavista and Yahoo both of the search engines give mix results in both the queries which show that Yahoo and Altavista are not able to differentiate between noun and verb sense of the word 'conduct'.

3) Phonetic Tolerance: There are numerous words which are phonetically equivalent but in writing they appear distinctively. Search engines should be capable of retrieving the results against any phonetically equivalent word of a keyword entered by the user. Indian languages alphabet contains many characters which give same sound i.e they are phonetically equivalent. These characters are used interchangeably in many modern Indian languages.e.g. झंडा, झँडा, झण्डा and झन्डा are phonetically equivalent words.

We had observed that Google and Guruji are not phonetic tolerant whereas Raftaar is not able to identify different phonetic words. The apparent variations are found in the relevancy of results on the variation of the phonetic keywords. English language is not at all phonetic in nature. So such problem does not arise with the English language based search engines.

B) Other factors

Besides the above mentioned factors in context of morphological structure of Hindi language. There are other critical challenges for the Hindi language and English language searching on the web, which are as follows-

1) Ambiguous keywords: Many words are polysemous in nature. Finding the correct sense of the words in the given context is an intricate task. Various researchers Eric Brill [5], Argaw [6], Navigili [7] and Christopher Stokoe and John Tait [8] and many others have justified the role of Word Sense Disambiguation in the improvement of performance of web searching. Ambiguous keywords deflate the relevancy of the results. The example mentioned below shows this aspect very clearly-

Query#1 सोना और स्वास्थ्य {Sona aur Swasthya} {Sleep and Health}

Query#2 गुलाब की कलम {Gulaab ki Kalam} (Rose Branch)

Query#3 दण्ड बैठक {Dand Baithak} (Dand Exercise)

Query#4 अंक विज्ञान {Ank Vigyan} (Numerology}

Query#5 कर्ण प्रिय संगीत {Karn Priy Sangeet} (Melodious Music)

TABLE III. PRECISION OF GOOGLE, RAFTAAR AND GURUJI IN CONTEXT OF SENSE AMBIGUITY FOR HINDI LANGUAGE

Query#	Precision on Google	Precision on Raftaar	Precision on Guruji
1	0.23	0.20	0.10
2	0.40	0.50	0.20
3	0.44	0.30	0.30
4	0.50	0.40	0.20
5	0.40	0.40	0.10

English language also faces the similar problem of ambiguous keywords like 'bank', 'bat' etc. We had selected the five sample queries from the complete set of queries which are as follows-
Query#1 built a bat house
Query#2 Capital tours
Query#3 Electrical current in Canada
Query#4 Free computer class.
Query#5 Clip art of light bulb

TABLE IV. PRECISION OF GOOGLE, ALTAVISTA AND YAHOO IN CONTEXT OF SENSE AMBIGUITY FOR ENGLISH LANGUAGE

Query#	Precision on Google	Precision on Altavista	Precision on Yahoo
1	0.56	0.64	0.67
2	0.45	0.43	0.56
3	0.71	0.56	0.45
4	0.64	0.39	0.57
5	0.46	0.46	0.44

It is quiet obvious from the results mentioned in Table 3 and 4 that in both languages, Search Engines are not capable enough to cope up with this problem. The results show the low precision values which justify that the performance of the search engines is affected by the sense ambiguity.

2) Word Synonyms: Every language has words and its synonyms. It is observed while working on English and Hindi language search engines that any word can express a myriad of implications, connotations, and attitudes in addition to its basic 'dictionary' meaning. Choosing the right word can be difficult for people, as well as for the Information Retrieval System. It is seen that most of the times when we alter the keywords with

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

their synonyms the performance of the search engines varies.

Query #1.1 आरक्षण से फ़ायदा {Arakshan se faida} (Benefits of reservation)

Query #1.2 आरक्षण से लाभ {Arakshan se laabh} (Benefits of reservation)

Query #2.1 भारत के राज्य-{Bhaarat ke rajya} (States of India)

Query #2.2 भारत के प्रदेश {Bhaarat ke Pradesh} (States of India)

Query #3.1 गाँधी जी की मृत्यु {Gaandhi ji ki Mrityu} (Death of Gandhi ji)

Query #3.2 गाँधी जी का निधन {Gandhi ji ka Nidhan} {Death of Gandhi ji}

Query # 4.1 पहला अंतरिक्ष यान {Pehla Antriksh Yaan}
(First Space ship)

Query # 4.2 प्रथम अंतरिक्ष यान (Pratham Antriksh Yaan) (First Space ship)

TABLE V. PRECISION OF GOOGLE, RAFTAAR AND GURUJI IN CONTEXT OF SYNONYM PROBLEM FOR HINDI LANGUAGE

Query	Results for Google		Results for Raftaar		Results for Guruji	
	Doc.	Pre.	Doc.	Pre.	Doc.	Pre.
1.1	3720	.70	4489	.50	24	.30
1.2	189,000	.90	15,408	.80	1625	.90
2.1	3,590,000	.70	319,903	.10	29,038	.40
2.2	2,670,000	.80	309,138	.10	37,750	.30
3.1	44,000	.60	12,844	0	516	.30
3.2	22,200	.10	18,455	0	517	.10
4.1	27,300	.80	2349	.40	6285	.70
4.2	13,000	.20	1314	0	308	.60

The variation in results is found in Query#1 and Query#2 when alternative synonyms are used. In Query#1 'लाभ' in place of 'फ़ायदा' and in Query#2 'प्रदेश' in place of 'राज्य' changed the result set. It is justified that replacing some words with their synonyms some times improves the results against the query in any language.

The problem of synonyms lies in every language. English language also faces the similar problem. Changing the keywords of the query with its synonyms in general vary the relevancy level of the results. From the complete set of 50 queries 5 queries are used to justify the impact of varying synonyms on the web search results in English language.

Query# 1.1 bed sharing with children
Query# 1.2 bed sharing with kids
Query# 2.1 school bus safety
Query# 2.2 school bus security
Query# 3.1 aircraft safety act of 2000
Query# 3.2 aircraft protection act of 2000
Query# 4.1 Freedom of information act forms.
Query# 4.2 Liberty of information act forms.

TABLE VI. PRECISION OF GOOGLE, RAFTAAR AND GURUJI IN CONTEXT OF SYNONYM PROBLEM FOR ENGLISH LANGUAGE

Query #	Results for (Google)		Results for (Altavista)		Results for (Yahoo)	
	Doc. Retrieved	Precision	Doc. Retrieved	Precision	Doc. Retrieved	Precision
1.1	1,070,000	0.44	82,200,000	0.65	77,300,000	0.47
1.2	1,180,000	0.56	68,500,000	0.48	62,900,000	0.43
2.1	9,460,000	0.67	47,400,000	0.45	48,700,000	0.49
2.2	12,500,000	0.68	51,400,000	0.57	48,000,000	0.67
3.1	231,000	0.68	6,480,000	0.35	39,200	0.68
3.2	420,000	0.66	6,480,000	0.36	77,300	0.65
4.1	20,600,000	0.47	62,900,000	0.48	725,000	0.57
4.2	908,000	0.39	19,200,000	0.35	253000	0.47

From the results mentioned in Table V and VI it is evident that the variation of query terms with their synonyms varies the precision level of the results.

IV. DISCUSSION

We have done comprehensive comparison of the performance of the English language and Hindi language based search engines in respect to their morphological structure and also other factors. On comparing the results of the query sample set it is concluded that when the query is given in its root form it returns into the maximization of results in Hindi language but only sometimes in English language. But in both the cases it is crystal clear from the results in Table 7 that the precision values are better when the key terms of the queries are in their root form.

From the results it is quiet evident that Google indexes the keyword in its root form for Hindi language but not for the English language. It is capable of listing the documents consisting of all

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

morphological variants of the keywords which justifies that the Google do not require stemmer for Hindi language web information retrieval. The other two search engines Raftaar and Guruji are not listing all the morphological variants of the query and hence they entail to have stemmer. For English language Web IR Altavista and Yahoo also show some improvement of results when query terms are used in their root forms. The results in Table VII justify that we do require converting the keywords in their root form to improve the relevancy of the results.

All the three search engines are not able to disambiguate the POS as mentioned in section 3. For English language Web IR only Google is capable of disambiguating POS whereas Altavista and Yahoo are not able to do that. Only Raftaar is upto some extent Phonetic tolerant. Google and Guruji are not phonetic tolerant for Hindi language. English language does not face the problem of phonetic terms.

None of the three (Google, Raftaar and Guruji) are attuned with ambiguity problem. Even English language search engines are also not attuned with ambiguity problem. On comparing the performance of the English and Hindi language search engines we found that precision values of English language is somewhat better than the Hindi language though the affect of ambiguity is quite visible on the performance of the Hindi as well as English language search engines. The Search Engines performance is degraded because of the sense ambiguity problem in Hindi as well as English language.

Sense / Synonym management is one of the challenges mentioned. It is marked that all the three search engines do not implement sense disambiguation or synonym management. Sense disambiguation improves the relevancy of results in web searching, as various researchers have successfully justified the application of WSD and synonym management in web searching.

It is apparent from the results that relevancy of the results retrieved by the search engines is dependent on the morphological structure and also, senses and synonyms in English as well as Hindi language.

An overall comparison we conclude that the performance of the English and Hindi language search engines affected by their morphology to some extent. Since the English language search engines are grown-up enough so the percentage of this affect is less in English language search engines as compared to Hindi language search engines.

V. CONCLUSION

We have compared the performance of the two language search engines i.e. English and Hindi in the light of their morphological structure and other factors also.

Our results conclude that the performance of the search engines is quiet affected by the morphological issues as well as sense ambiguity and synonym problems. This affect is much obvious in the Hindi language search engines in comparison to the English language.

The morphological structure of Hindi language is more critical in comparison to the English language. This is the reason the performance of the Hindi language is more affected by such issues.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

TABLE VII. PRECISION OF GOOGLE, RAFTAAR AND GURUJI IN CONTEXT OF SYNONYM PROBLEM FOR ENGLISH LANGUAGE

Parameters	**Hindi language Search Engines Performance**			**English language Search Engines Performance**		
	Google	Raftaar	Guruji	Google	Altavista	Yahoo
Root Word						
With Root Word	0.80	0.57	0.40	0.59	0.53	0.55
Without Root Word	0.77	0.30	0.37	0.54	0.56	0.53
Sense Ambiguity	0.39	0.36	0.18	0.56	0.50	0.54
Synonym Management	Variation in Precision	Variation in Precision	Variation in Precision	Variation in Precision	Variation in Precision	Variation in Precision
POS Disambiguation	Do not support	Do not support	Do not support	Support to some extent	Do not Support	Do not Support
Phonetic Tolerance	Not Phonetic Tolerant	Not Phonetic Tolerant	To some extent Phonetic Tolerant	Not a phonetic Language		

REFERENCES

[1] P. Bhattacharya, S. Singh, K. Gupta, and M. Srivastava, "Morphological richness offsets resource demand- experiences in constructing a POS tagger for Hindi" in Proceedings of COLING, 06, Sydney, Australia, pp.- 779-786, 2006

[2] http://www.google.com

[3] http://www.raftaar.com

[4] http://www.guruji.com

[5] E. Bril and S. Vassilvitskii, "Using WebGraph Distance for Relevance Feedback in Web Search" in Proceedings of SIGIR'06, Seattle, Washington, USA, pp. -147-153, 2006

[6] A. A. Argaw, "Amharic-English Information Retrieval with Pseudo Relevance Feedback", in Proceedings of 8th Workshop of the Cross-Language Evaluation Forum, CLEF 2007, Budapest, Hungary, pp. 119-126, 2007

[7] R. Navigili and P. Velardi, "Structural Semantic Interconnection: a knowledge-based approach to Word Sense Disambiguation", in Journal of Pattern Analysis and Machine Intelligence, Volume 27, Issue 7, pp. 1075 – 1086, July 2005

[8] C. Stokoe and J. Tait, "Towards a Sense Based Document Representation for Internet Information Retrieval", in Proceedings of SIGIR'03, July 28- August 1, Toronto, Canada, pp. 791-795, 2003

Comparison of Traffic in Manhattan Street Network in NS2

Ravinder Bahl *(Author[1])*
Information and Technology
M.M.E.C
Muallana, Ambala, Haryana, India
ravindra_ibm@yahoo.com

Rakesh Kumar *(Author[2])*
Information and Technology
M.M.E.C
Muallana, Ambala, Haryana, India
raakeshdhiman@gmail.com

Rakesh Sambyal*(Author[3])*
Information and Technology
MBS College of Engineering and Technology
Babliana, Jammu, Jammu and Kashmir,India
rakeshsambyal@rediffmail.com

Abstract— **The paper presents the Comparison analysis of traffic in Manhattan Street Network (MSN). The behaviour of Manhattan Street Network for constant bit rate (CBR) and exponential traffic sources is demonstrated. The results are produced using NS2 simulator. It is concluded that the performance in multipath networks like MSN can be improved by taking account of appropriate buffering for each traffic source at the link node.**

Keywords— Comparison Analysis, ***Traffic, Constant bit rate (CBR), Exponential, MSN, Buffering.***

I. INTRODUCTION

The selection of buffer is very important as it helps in reducing the congestion and drop in packets .It helps in balancing the load across the multipath and multihop networks like Manhattan Street Network (MSN).The traffic sources like constant bit rate (CBR) and exponential offer different type of load in the network. In the large networks like Manhattan Street Network the traffic is routed form two input links to two outgoing links having equal bandwidth and delay. The packet is stored in buffers upon arrival and is deflected on the other. It cannot be accommodated at a later stage even if the buffer allows admission to new arrivals at that stage. In voice network and private line data networks the capacity planning is very simple and straight forward with problem of congestion and packet drops. For the selection of buffer capacities, the growing traffic volume is to be calculated. As the traffic increases so is the demand for more bandwidth. The appropriate size and type buffer selection can reduce the growing demand for more bandwidth. In this paper, attempt has been made make provision for reducing congestion and packet drops in different traffic sources like constant bit rate (CBR) and exponential. In order to reduce the congestion and packet drop , the drop tail queue is provided at each input node which also help in reducing the deflections. Further regardless of the destination, the buffering structure is so designed to store the packets and deflection occurs only when all the buffer slots are full. The distance vector routing(RIP) and link state (OSPF) play a very important role in today's internet [1].The simulation comparison of buffers in OSPF and RIP algorithm conclude that the link state routing performs better than distance vector routing in case of large networks like Manhattan Street Network topology where large volume of traffic flows form source to destination.

II. BACKGROUND OF ROUTING ALGORITHM

In case there are more than two paths from source to the destination. The path selected between source and destination represents the outgoing link to be used. The routing algorithms based on cost metrics [5] helps in calculating the path between the nodes. The cost is defined in terms of number of hops or bandwidth, number of links, distance. This depends upon the metric supported by different of buffer capacities. Many different cost metrics [5] can be used to judge the shortest path between the source and destination.

III. CONCEPT OF MULTIPATH ROUTING ALGORITHM

In multipath networks where data transactions in large volume take place from source to destination, the path with the optimum cost is to be selected to route the data. The selection of improper path leads to over utilization of the network resources resulting increasing delays and congestion. core networks where there is a huge amount of data transactions and there are more than one equal cost route possible from a source node to destination node with large volume of traffic, the multipath routing algorithm [6][7][9] may be used, which helps in improving the available resources utilization and helps in reducing congestion and packet drops and thus helps in shaping the traffic between equal cost multi paths. The links utilization [2] can be improved. By having improper queues which in turn implies increasing delays, so sort of trade off is required for selection of appropriate queue. The simulation has

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

proved that appropriate size queue for particular traffic source results in better resource utilization .Results help in deciding the size of queue at the link for better performance.
The selection of appropriate size queue avoids congestion at a particular node including cross-traffic as given below.

(A) Departure of traffic from a node should be equal to the traffic arriving at that node.
(B) It is essential in case of multipath networks that the load should be balanced in such a way so that none of the outgoing link of a router is over utilized.
(C) The appropriate queue with optimum storage capacity may be maintained at the bottleneck links for better performance.

IV. MANHATTAN STREET NETWORK (MSN) TOPOLOGY

(A) The topology used in the paper for the purpose of demonstration has 16 nodes and 32 links of equal bandwidth and delay.
(B) The two traffic sources, constant bit rate and exponential are used at the link node one at a time.

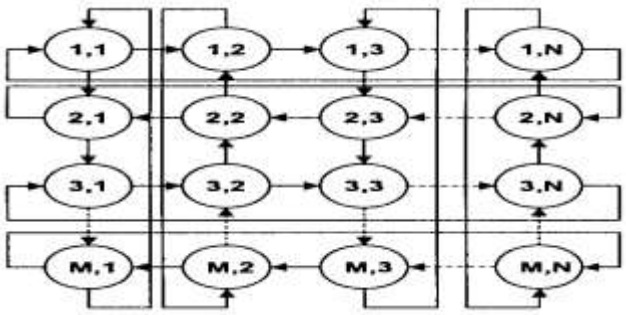


Fig.(1) : Experimental Topology (MxN) , Where (M=4,N=4)

V. SIMULATION RESULTS

The simulation study was performed and analysis were drawn using Network Simulator (NS2) [8]. The Simulation results are evaluated for in different buffering capacities required for each traffic source using topology of Fig. (1) on case to case basis.

VI. BOTTLENECK LINK CASES

The traffic from node n1 to n16 was observed where the link bandwidth was set to 1Mbps, delay of 10 ms, and packet size of 64 bytes. Initially two constant bite rate traffic sources and later two exponential traffic sources with interval of 0.005 seconds and drop tail queue was introduced at a time. Initially queue size was set to no packet and later was changed to the capacity of more packets and performance was analyzed. The simulation was run for 10 seconds and traffic was introduced for 5 Seconds and the Following Results were obtained.

TABLE I : SHOWING SIMULATION RESULT WITH QUEUE SIZE OF NO PACKET.

Packet Size(bytes)	Traffic Type	Total Packet	Non Dropped Packets	Dropped Packets
64	CBR	2002	2002	0
64	Exponential	3110	3110	0

TABLE I : SHOWING SIMULATION RESULT WITH QUEUE SIZE OF 1 PACKET.

Packet Size(bytes)	Traffic Type	Total Packet	Non Dropped Packets	Dropped Packets
64	CBR	2002	0	2002
64	Exponential	3110	0	3110

TABLE III: SHOWING SIMULATION RESULT WITH QUEUE SIZE OF 2 PACKETS.

Packet Size(bytes)	Traffic Type	Total Packet	Non Dropped Packets	Dropped Packets
64	CBR	2002	2002	0
64	Exponential	3110	3110	0

All links set to type simplex having bandwidth (1Mbps), delay (10ms) the offered load was observed between n1 to n16 with two constant bit rate and two exponential traffic sources starting at n1 and n16 as destination. The drop tail queue was attached between n1 and n16.

VII. PERFORMANCE GRAPHS

The performance graph for bottleneck link, when one drop tail queue with FIFO (First in First out) discipline is maintained at the bottleneck link with two different traffic sources as shown in Fig.(2). , Fig. (3), Fig. (4) and Fig. (5) Respectively.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

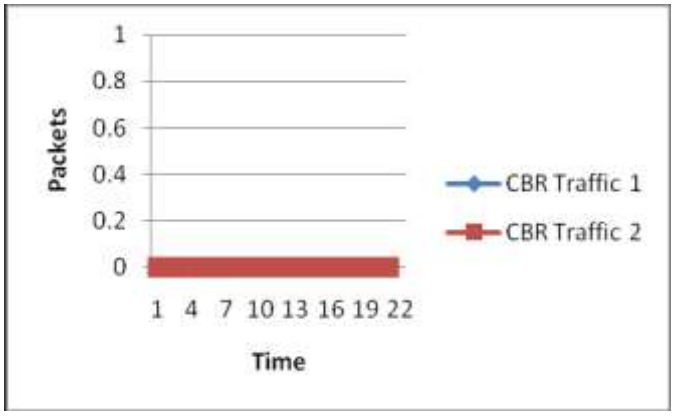


Fig. (2): CBR Traffic in Packets with Drop Tail Queue of size 0

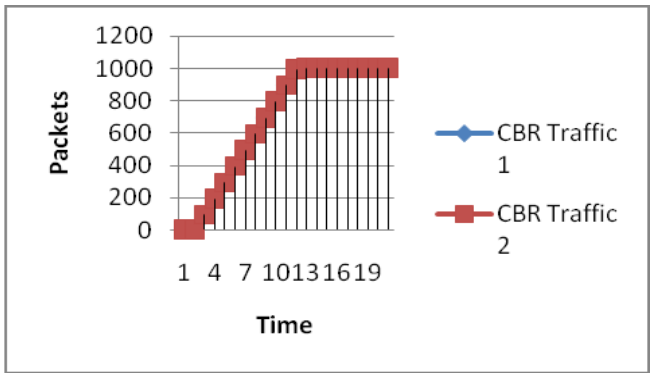


Fig. (3): CBR Traffic in Packets with Drop Tail Queue of size 2

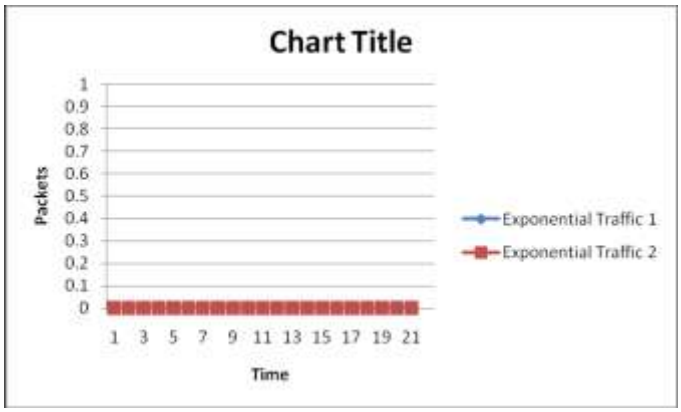


Fig. (4): Exponential Traffic in Packets with Drop Tail Queue of size 0

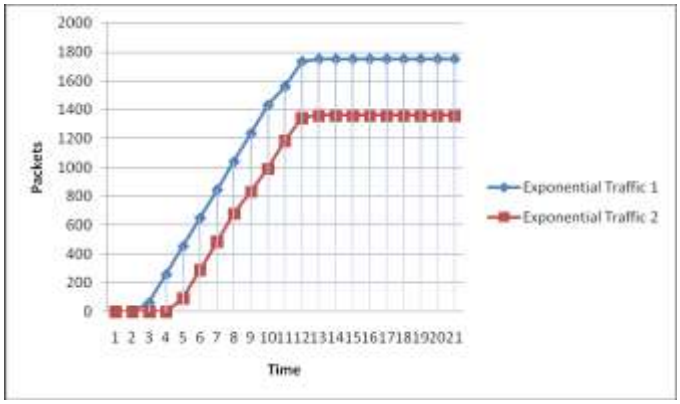


Fig. (5): Exponential Traffic in Packets with Drop Tail Queue of size 2

The simulation result helps in deciding the type of buffer at the bottleneck links. Case study of buffer types for a particular case is explained. The decision can be made based on the performance graph that the appropriate buffer is implemented at the bottleneck links for a particular case such that the congestion in the network may be reduced and performance may be improved.

VIII. CONCLUSION

The demonstration of buffer selection at bottleneck links in the Manhattan Street Network Topology is done.

The paper demonstrated with different types of buffer , packet size of 64 bytes and two constant bit rate (CBR) traffic sources starting form node n1 for destination n16 the following effects

Case (1) Simulation of multipath network topology (MSN) with drop tail node queue having storage capacity of 0 packets with two constant bit rate traffic Sources led to total packet drop.
Case (2) Simulation of multipath network topology (MSN) with drop tail node queue having storage capacity of 0 packets with two exponential traffic Sources led to total packet drop.
Case (3) Simulation of multipath network topology (MSN) with drop tail node queue having storage capacity of one packet with two constant bit rate traffic Sources led to total packet drop
Case (4) Simulation of multipath network topology (MSN) with drop tail node queue having storage capacity of one packet with two constant bit rate traffic Sources led to total packet drop
Case (5) Further it was concluded that Simulation of multipath network topology (MSN) with drop tail node queue having storage capacity of two packet with either two constant bit rate or exponential traffic Sources led to no packet drop

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

IX. REFERENCES

[1] Routing Basics,http:// www.cisco.com/ univercd / cc / td /doc/cisintwk/ ito_doc/ routing.htm.

[2] James Irvine, David Harle, "Data Communication and Networks", John Wiley & Sons Ltd.,New York, USA,2002.

[3] William Stallings ,"Data and Computer Communications", PHI Pvt. Ltd. N.Delhi, 7th Edition,2003.

[4] Alberto Leon-Garcia, Indra Widjaja,:"Communication Networks, Fundamental Concepts and Key Architectures",Tata McGraw-Hill Publishing Company Ltd.,N.Delhi,2nd Edition, 2005.

[5] Brian Hill , "The Complete Reference,CISCO", Tata McGraw-Hill Publishing Company Ltd.,N.Delhi, 3rd reprint 2004.

[6] Johnny Chen, Peter Druschel, Devika Subramanian, "An Efficient Multipath Forwarding Method",Proceedings of IEEE INFOCOM , San Francisco, CA,March 1998.

[7] Israel Cidon, Raphael Rom, Yuval Shavitt, "Analysis of Multi-path Routing", IEEE/ACM Transanctions on Networking, Vol.7,No.6,Dec.1999.

[8] The Network Simulator ns-2, http://www.isi.edu/nsnam/ns/ns-documentation.

[9] Ivan Gojmerac, Thomas Ziegler, Fabio Ricciato, Peter Reichl, "Adaptive Multipath Routing for Dynamic Traffic Engineering", IEEE GLOBECOM-2003,http:// userver.ftw.at /~reichl /publications /GLOBECOM03.pdf.

An Evolving Order Regularized Affine Projection Algorithm, Suitable For Echo Cancellation

Shifali Srivastava
Electronics Deptt.
JIIT,
Noida, India
shifalihbti2004@gmail.com

M.C.Srivastava
Electronics Deptt.
JIIT
Noida, India
m.c.srivastava@jiit.co.in

***Abstract*— In this paper, a regularized Affine Projection algorithm with Evolving Order is proposed. This algorithm automatically determines its projection order, derived in the context of acoustic echo cancellation (AEC). The simulation results indicate that the proposed algorithm yields better performance with small steady state error as compared to existing evolving order affine projection algorithm (APA) and has fast convergence speed.**

Keywords—Acoustic echo cancellation (AEC),affine projection algorithm (APA),evolving order affine projection algorithm (EO-APA), Evolving order regularized affine projection algorithm (EO-RAPA),double talk(DT),echo path change(EPC).

I. INTRODUCTION

In acoustic echo cancellation (AEC) contexts the basic approach is to build a model of the impulse response of the echo path using an adaptive filter, which provides replica of the echo at its output [1]. The adaptive filter output is subtracted from the microphone signal to cancel the echo. Several challenges are associated with AEC applications. Firstly, the echo path can be extremely long and it may rapidly change. Secondly, the background noise that appears at the near-end side can be strong and non-stationary in nature. Further, the involved signals (i.e., speech) are non-stationary and highly correlated.

For echo cancellation several adaptive algorithms [1], [2] have been applied. The normalized least-mean square (NLMS) algorithm and the affine projection algorithm (APA) are preferred due to their simplicity and robustness. The affine projection algorithm (APA) [4], [5] updates the weights based on the last input vectors. The convergence speed of APA for correlated input signal is improved by employing an updating-projection scheme of an adaptive filter on a P-dimensional data-related subspace, but the convergence speed of APA decreases, in the presence of noise.

Recently, EO-APA has been proposed to improve performance in the presence of noise, with fast convergence speed and small steady state error by varying the number of input vectors [7]. An evolutionary method is employed in EO-APA to determine necessary number of input vectors. In this algorithm order of vector increases or decreases from the previous one by comparing the output error with the threshold involving the information of the steady-state mean-squared error (MSE) [8].

The convergence speed of EO-APA gets degraded during unvoiced speech signal and/or during silences when signal value is either zero or close to zero. This paper proposes solution of such problems by employing an evolving order regularized Affine Projection Algorithm (EO-RAPA). In the proposed algorithm, regularization is obtained by adding a regularization factor matrix prior to taking inverse of the correlation matrix. With the suitable choice of regularization factor, the performance of the algorithm can be improved with decreased computational complexity.

A major concern associated with the behavior of the algorithm during double-talk, that may affect the overall performance of proposed algorithm, is also addressed in the proposed scheme. The standard procedure can be used for double talk detection in order to slow down or completely halt the adaptation process during double-talk periods.
Several algorithms have been proposed for detection of double talk (DT) [9]–[11]. The simplest double talk detection algorithm is the well-known Geigel DTD [11], which provides a low-complexity solution. Since it is not efficient to distinguish between echo path change and double talk, a modified version of it is proposed in this paper.

This paper is organized as follows. Section II describes evolving order APA (EO-APA). In section III, we propose a regularized version of evolving order APA (EO-RAPA) for AEC and performance improvement during double talk. The experimental results which illustrate the convergence performance of the proposed algorithm are discussed in section IV. Finally, conclusions are presented in Section V.

II. EO-APA

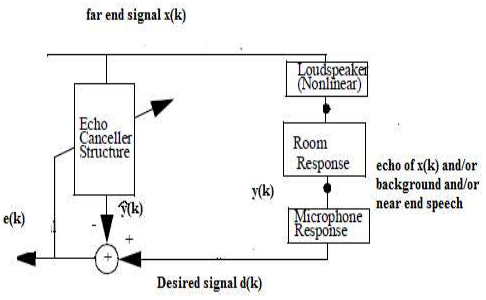


Figure 1 Acoustic echo canceller structure

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

A general acoustic echo canceller structure is depicted in Fig. 1.The aim of this structure is to model the acoustic echo path (i.e. room impulse response) using an adaptive filter. As shown in figure the far end signal *x(k*), on being processed by the room impulse response, produces echo signal. This echo signal is picked up by the microphone and/or the near end signal *v(k)* and/or background noise *n(k)* to produce a desired signal *d(k).*The far end signal that also acts as the input to the adaptive filter produces a output $\hat{y}(k)$, the replica of echo signal.

The error signal $e(k) = d_k - \hat{y}(k)$, is used for weight adaptation of the filter. With the far end signal denoted by $\boldsymbol{X_k}$ in the matrix form, the desired $\boldsymbol{d}_k$ may be expressed as:

$$\mathbf{d}_k = \boldsymbol{W}^T_{opt}\boldsymbol{X_k} + n_k \tag{1}$$

where $\boldsymbol{X_k} = [x_k , \ x_{k-1}, \ . \ . \ ., \ x_{k-L+1}]^T$,the subscript k represents the time index, $\boldsymbol{W}_{opt}$ is an unknown L × 1 weight adaptation column vector and n_k is a zero-mean Gaussian independent sequence with variance of σ_n^2 .

An estimate $\mathbf{W}_k$ for $\boldsymbol{W}_{opt}$ at iteration k may be computed as follows [4] by using an affine projection algorithm (APA):

$$\mathbf{W}_k = \mathbf{W}_{k-1} + \mu\mathbf{U}_k^H(\mathbf{U}_k\mathbf{U}_k^H)^{-1}\mathbf{e}_k \tag{2}$$

where μ is a constant step size .The error signal $\mathbf{e}_k$ may therefore be expressed as:

$$\mathbf{e}_k = \mathbf{d}_k - \mathbf{U}_k\mathbf{W}_k^T \tag{3}$$

where

$$\mathbf{d}_k = [d_k , d_{k-1} , \ldots , d_{k-P+1}]^T \tag{4}$$

The matrix in (3) $\mathbf{U}_k$ is the collection of the P most recent input vectors $[\boldsymbol{X}_k, X_{k-1}, \cdots , \boldsymbol{X}_{k-P+1}]^T$.

The order of the APA is defined by the projection order P, ,the number of the input vectors used to determine $\mathbf{U_k}$, that should be less than or equal to filter length L [1]-[3]. In EO-APA the projection order P, at any iteration that varies in accordance with adaptation, may be represented as [7]:

$$P_k = f(P_{k-1}, \mathbf{e}_k^2) \tag{5}$$

Since the number of input vectors at any iteration depends on the output error and the previous number of input vectors, the projection order adjusts itself if error exceeds a particular threshold that can be determined by approximating steady-state MSE.

Following Shin and Sayed [8], for each iteration the threshold error may be written as:

$$\eta_k = \frac{\sigma_n^2\{\mu(P-1)+2\}}{2-\mu} = \varepsilon(P) \tag{6}$$

If square of the error signal $e^2(k)$ is smaller than $\varepsilon(P_{k-1})$, projection order P_k at the k^{th} iteration in EO-APA should be reduced by one from P_{k-1} (its previous value), for smaller steady state error. Whereas, P_k should be increased by one from P_{k-1} when $e^2(k)$ is larger than, $\varepsilon(P_{k-1}+1)$ for faster convergence speed. Therefore, the upper and lower thresholds η_k and θ_k at k^{th} iteration respectively may be expressed for EO-APA can be expressed as:

$$\eta_k = \varepsilon(P_{k-1}+1) = \frac{\sigma_k^2\{\mu(P_{k-1}+2)}{2-\mu} \tag{7}$$

and

$$\theta_k = \varepsilon(P_{k-1}) = \frac{\sigma_k^2\{\mu(P_{k-1}-1)+2\}}{2-\mu} \tag{8}$$

With these thresholds bound the projection order at any iteration may be determined as:

$$P_k = \begin{cases} Min\{P_{k-1}+1, P_{\max}\} & \text{if } \eta_k < e^2(k) \\ P_{k-1} & \text{if } \theta_k < e_k^2(k) \le \eta_k \\ Max\{P_{k-1}-1, 1\} & \text{if } e_k^2 \le \theta_k \end{cases} \tag{9}$$

The upper threshold controls the projection order of the EO-APA to track the increased output error due to variation in environment. Similarly the lower threshold acts as a switching point to decrease the projection order. Thus EO-APA is expected to provide fast convergence speed with decrease in steady-state error.

III. EO-RAPA

H. Rey, L. Rey Vega, S. Tressens, and J. Benesty in [13], employed a variable explicit regularization factor in their work on APA. They suggested an optimal value of regularization factor (δ) chosen such that minimizing difference of weight error vector of the consecutive iterations resulted in lower steady state error. Such a choice maximizes speed of convergence and minimizes steady-state mismatch. Under simplifying assumption they define δ_k as:

$$\delta_k = \frac{P.\sigma_v^2.\sigma_x^2.L}{E\left[\left\|e_k^2\right\|\right] - P\sigma_v^2} \tag{10}$$

In the proposed work a regularization factor has been added in the EO-APA. During unvoiced speech or during silences, the signal level is either very low or zero and therefore inverse of $(\mathbf{U}_k\mathbf{U}^H_k)$ gets ill conditioned in weight updation in (2). Further, with the increase in the projection order of the EO-APA, computational effort for determining inverse of the matrix increases. To avoid these problems a regularization factor is added before taking inverse in the proposed algorithm. The regularization factor also helps in suppressing the effect of noise, which may be the background noise or/and the near-end speech corrupting the output of the echo path. The proposed algorithm referred as evolving order regularized APA (EO-RAPA), may be expressed as:

$$\mathbf{W}_k = \mathbf{W}_{k-1} + \mu\mathbf{U}^H_{k,P_k}(\mathbf{U}_{k,P_k}\mathbf{U}^H_{k,P_k} + \delta\mathbf{I})^{-1}\mathbf{e}_{k,P_k} \tag{11}$$

The identity matrix $\mathbf{I}= P_k*P_k$ and δ is the regularization factor.

Performance improvement during Double Talk:

When the speech signal *v (k)* is zero and the near-end noise *n (k)* is assumed to be insignificant, weights of the adaptive filter would converge to successfully cancel the echo, and successfully cancel the echo. However, when both

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

v(k) and *x(k)* are not zero, that is in double-talk (DT) situation, the near end speech *v(k),* acts as an uncorrelated noise to the adaptive algorithm, and may allow excessive uncancelled echo to pass. Solution to this problem is to slow down or completely stop the filter adaptation when the presence of the near-end speech is detected.

A simple approach to detect DT is the well-known Geigel algorithm [11]. The Geigel algorithm compares the magnitude of the near-end received signal *d (k)* with the maximum magnitude of L most recent samples of the far-end signal *x(k),* where L is the adaptive filter's length. The Geigel algorithm computes its detection variable and makes decision. If detection variable is larger than the threshold , DT is declared otherwise not. However, for an AEC environment, the echo path characteristics are time varying. Therefore, for the time-varying echo path, DT can be falsely detected by the Geigel algorithm when a change of the echo path occurs. As a result, the adaptive filter stops updating the coefficients when the coefficient update is actually needed.

In the proposed algorithm EO-RAPA a detection scheme is presented that detects and distinguishes between echo path change and double talk. Simulation results show that this detection method provides adequately reliable performance with lower complexity. This method is based on computation of detection variable ξ at any iteration:

$$\xi = \frac{\text{mean power of microphone signal } \mathbf{d}(k)}{\text{mean power of far end signal } \mathbf{X}(k)} \qquad (13)$$

If $\xi >$ threshold T_h and at the same time projection order is above the certain value say P_{th}, double talk is detected. Since projection order is directly related to the output error that corresponds to the cross correlation between microphone signal and far-end signal. Thus, our detection method is power and correlation based method.

In the case of EPC condition, projection order is above a certain value P_{th} but $\xi <$ threshold T_h. We therefore employ two thresholds one on ξ and one on projection order, to reduce the probability of false detection of DT or EPC conditions. Flow chart of DT detection scheme is shown in Fig. 2.

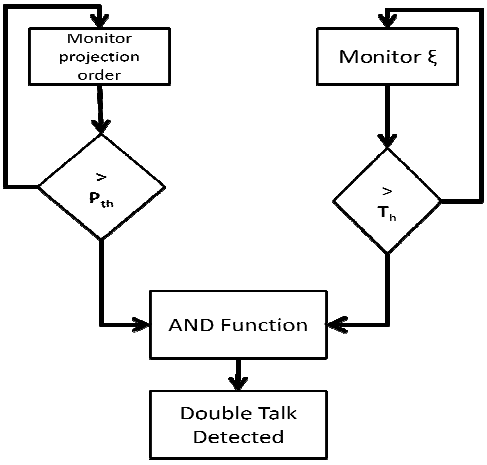


Figure 2 Flow chart of DT detection scheme

IV. SIMULATION RESULTS

The performance of the proposed algorithm is verified by carrying out experiments in context of echo cancellation with speech input sampled at 8 KHz and echo path length of 512.

ERLE, an important parameter to measure of convergence speed and misadjustment that may be expressed in dB by (13), is employed for performance evaluation.

$$ERLE = 10 * \log \frac{E\left[\sigma_e^2\right]}{E\left[\sigma_d^2\right]} \qquad (13)$$

where σ_d^2 is the power of the microphone signal and σ_e^2 is the power of the residual echo.

In Fig. 3, by ensemble averaging over 20 independent speech samples, ERLE plot is obtained for proposed algorithm with variable regularization factor, EO- APA and classical APA for different projection orders (*P*=8 and 16), SNR=30 dB. As shown by simulation results the performance of proposed algorithm is better in terms of convergence speed and low steady state error.

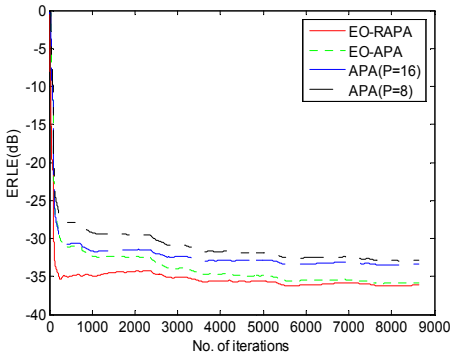


Figure 3 ERLE plot of proposed algorithm, EO- RAPA and classical APA (echo path length =512, speech input sampled at 8000 Hz, step-size=0.5)

Fig. 4 Shows ERLE plot for different values of fixed regularized factor (δ). Simulation results show the practical justification that as δ increases, steady state mismatch reduces but at the cost of lower convergence speed. For $\delta = 0.5 * \sigma_x^2$, convergence speed is faster but steady state error increases. On the other hand for $\delta = 500 * \sigma_x^2$ slower convergence speed is obtained with lower steady state error.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

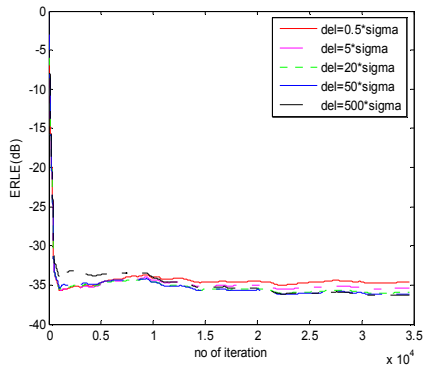


Figure 4 ERLE plot of proposed algorithm with different values of fixed regularization factor (where del is regularization factor and sigma is mean power of far end signal), SNR=30dB

Double talk scenario is depicted in Fig.5, ERLE plot of proposed algorithm when DT occurs from 19,000 to 21,000 with proposed double talk detection (DTD) scheme and without DTD scheme. Simulation results indicate that deterioration in performance during double talk (DT) is improved by employing proposed DTD scheme.

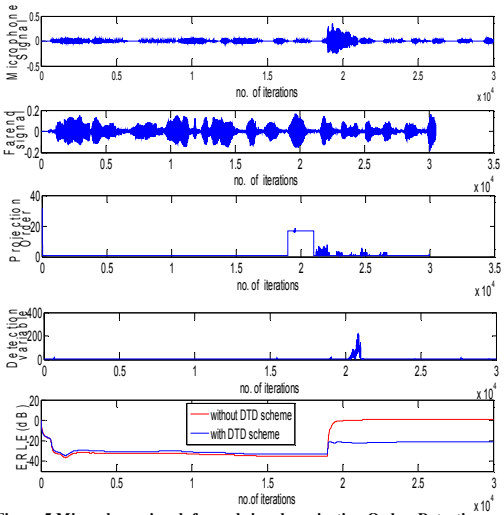


Figure 5 Microphone signal, far end signal, projection Order, Detection variable signal and ERLE plot of proposed algorithm when DT occurs between 19,000 to 21,000

Fig. 6 shows plots of different signals and it can easily seen from simulation results that proposed DTD scheme detect and distinguish EPC and DT.

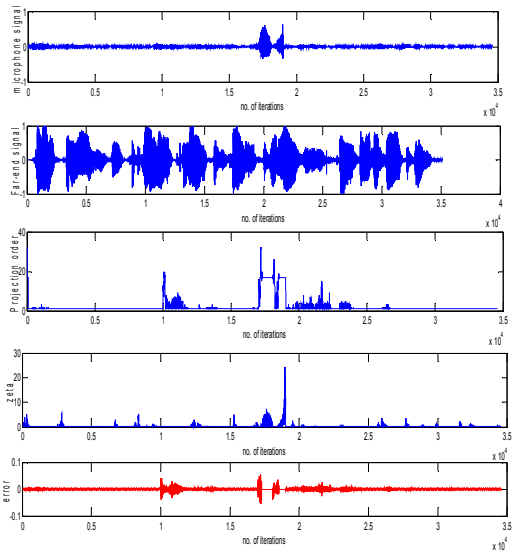


Figure 6 Microphone signal, Far end signal, projection Order, Detection variable signal (Zeta) and error plot of proposed algorithm when EPC occurs at iteration 10,000 and DT occurs between 17,000 to19,000 with DTD scheme, SNR=20dB

Fig 7 shows the performance of proposed algorithm by employing DT detection scheme when EPC and DT occurs at different iterations.

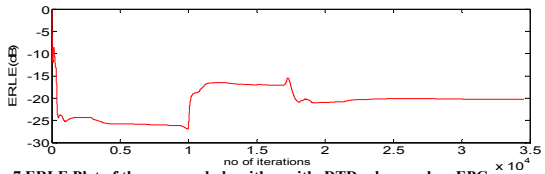


Figure 7 ERLE Plot of the proposed algorithm with DTD scheme when EPC occurs at iteration 10,000 and DT occurs between iterations 17,000 to 19,000, SNR=20dB

V. CONCLUSIONS

In this paper an evolving order regularized affine projection algorithm (EO-RAPA) has been evolved that is suitable for AEC applications. The performance of the proposed algorithm is verified by carrying out experiments in context of echo cancellation with speech input sampled at 8KHz. Simulation results indicate that proposed algorithm has faster convergence speed and lower steady state mismatch compared to existing EO-APA. Further the proposed algorithm provides an improvement in the performance during double talk by employing a new DTD scheme.

REFERENCES:

1. S. Haykin, Adaptive Filter Theory, 4th ed. Upper Saddle River, NJ: Prentice Hall, 2002.
2. A.H.Sayed, Fundamentals of Adaptive Filtering. New York, Wiley, 2003.
3. Simon S. Haykin, Bernard Widrow, Least-mean-square adaptive filters, Wiley ;2003.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

4. S.L. Gay and S. Tavathia, “The Fast Affine Projection Algorithm,” Proc. 1995 IEEE Int. Conf. Acoustics, Speech, and Signal Process, May 1995, vol. 5, pp. 3023–3026 .
5. K. Ozeki, T. Umeda, "An Adaptive Filtering Algorithm Using an Orthogonal Projection to an Affine Subspace. and Its Properties," Electronics and Communications in Japan, Vol.67-A, no. 5, pp. 19–27, 1984.
6. Hyun-Chool Shin, Ali H. Sayed, and Woo-Jin Song, “Variable Step-Size NLMS and Affine Projection Algorithms”, IEEE, Signal Processing Letters, Vol.11,No.2, Feb. 2004.
7. Seong-Eun Kim, Se-Jin Kong , Woo-Jin Song ,”An Affine Projection Algorithm With Evolving Order ”, IEEE Signal Processing Letter, Vol. 16, No.11, Nov 2009.
8. H.-C. Shin and A. H. Sayed, “Mean-square performance of a family of affine projection algorithms,” IEEE Trans. Signal Processing, vol. 52, no.1, pp. 90–102, Jan.2004.
9. J. Benesty, D.R. Morgan, and J. H. Cho, “A new class of doubletalk detectors based on cross-correlation,” IEEE Trans. Speech Audio Process., vol. 8, no. 2, pp. 168–172, Mar. 2000
10. Thien-An Vu, Heping Ding, and Martin Bouchard “A survey of double talk detection scheme for echo cancellation applications”
11. D.L. Duttweiller, “A twelve-channel digital echo canceller,” IEEE Trans. Comm., vol. 26, pp. 647-653, May 1978.
12. I.Yamada, K.Slavakis, and K.Yamada, “An efficient robust adaptive filtering algorithm based on parallel subgradient projection techniques,” IEEE Trans. Signal Process., vol. 50, no. 5, pp. 1091–1101,May 2002
13. H. Rey, L. Rey Vega, S. Tressens, and J. Benesty, “Variable explicit regularization in affine projection algorithm: Robustness issues and optimal choice,” IEEE Trans. Signal Process., vol. 55, no. 5, pp. 2096–2108, May 2007.

AUTHORS PROFILE

Shifali Srivastava received her B. Tech degree from H.B.T.I Kanpur, and now perusing M.Tech from JIIT, Noida. She has done projects on all optical networks. Her area of interest is signal processing and communications.

M.C. Srivastava received his B.E. degree from Roorkee University (now IIT Roorkee), M.Tech from Indian Institute of technology Mumbai and Ph. D form IIT Delhi in 1974.He was associated with I.T. BHU, Birla institute of Technology and Science Pilani, Birla institute of Technology Ranchi and ECE Dept. JIIT Sector 62 Noida. He has published around 60 research papers. His area of research is signal processing and communications. He was awarded with Maghnad Saha Award for his research paper.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Design and Implementation of Flexible Framework for Secure Wireless Sensor Network Applications

Inakota Trilok
Department of Computer Science & Engineering
National Institute of Technology
Warangal, India
itrilok@hotmail.com

Mahesh U.Patil
National Ubiquitous Computing Research Centre
Centre for Development of Advanced Computing
Hyderabad, India
maheshp@cdac.in

***Abstract*—Secure communications is an interesting and challenging research area in Wireless Sensor Networks (WSN) fundamentally because of the low power constraints and small memory footprints inherent in the technology. In this context, there are many hardware platforms like TelosB, MicaZ and Mote2 which implement a security layer in hardware, supporting multiple modes of operation like encryption, integrity or combinations of both. However, not all hardware platforms support hardware security which creates avenues of research in designing low power security algorithms in software. As with the development of security algorithms for WSN applications, there is an urgent requirement to create a unified approach for application developers by which they can integrate and use existing security algorithms thereby maintaining an abstraction from the intricacies of the algorithm.**

This paper introduces a flexible framework which implements a unified API to add new security algorithms to a security library suite. This library integrates existing security algorithms like TinySec, MiniSec etc. We also bring out the implementation of Advanced Encryption Standard (AES) in software supporting its various modes of operation. We have integrated this implementation with the unified framework and demonstrated its performance and our results. We compare our software AES implementation with the Hardware AES implementation, in all the supported mode settings.

Keywords-Wireless Sensor Networks; Mote; Link Layer Security; Network Layer Security; Hardware Level Security, Integrity; Encryption; Authentication.

I. INTRODUCTION

Wireless Sensor Networks (WSN) is a collection of distributed autonomous systems called sensor nodes that monitor and collect physical data for assessment and evaluation. These sensor nodes are very small in size, and are limited in resources like CPU, memory and network bandwidth. Moreover they are powered through small batteries. All these make wireless sensor networks vulnerable to security attacks and this is a crucial aspect of the sensor network. Much of security is application specific and in applications like physical intrusion detection or perimeter protection these are of utmost importance. As sensor nodes are powered through batteries, security techniques must ideally consume less energy. Ironically, some sensor networks need high security which leads to higher consumption of energy. In such cases, high-energy and rechargeable batteries are used. There are some contradictions between the communication energy cost and cryptographic cost for WSN (See [1, 2]). So, the security technique needed depends on the application that is to be deployed into the network. Thus, there needs to be a balance between the amount of security that can be provided to these networks and resources on the mote. This is different from conventional security solutions, since in WSN the security is tightly coupled to the application's need.

TinyOS [22] is a free and open source embedded operating system which is specifically designed for wireless sensor network application development. It follows a component based architecture which enables application developers to integrate their application requirements with existing network communication protocols. The application and operating system is bundled into a final image which is burnt onto the hardware thereby creating two tier architecture. Application developers should wire a customized network stack for which knowledge of low level details of each of the algorithms is required. These algorithms could spawn diverse areas like network communication, dissemination, time synchronization, and security, etc. Moreover, to modify the stack inorder to test performance, knowledge of the interfaces provided by alternate algorithms also have to be understood. This requirement imposes an additional burden on the developer.

There are number of security algorithms available in TinyOS communication protocol stack. Some of them like TinySec [3], SenSec [4] and CC2420/CC2430/CC2431 Radio AES [5] operate at the link layer, while other algorithms like MiniSec [6] exist at the network layer. As is seen above there is wide diversity in the implementation and detail of the algorithms. This renders migration from one security algorithm to another a point of bother for the application developer. A uniform access method for all security algorithms is desirable. The main contributions of this paper are:

- Introduction to an adaptive framework for WSN applications.
- A general purpose security library suite composed of existing security algorithms for popular sensor node hardware platforms.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

- An application developer's perspective in creating a custom network stacks specific to his application requirements.
- Support for existing versions of TinyOS providing the application developer an abstraction to low level implementation of the desired security algorithm chosen. The application developer is provided with a common API's for setting modes and updating keys for all available security algorithms.
- Implementation and integration of software AES and hardware AES with same mode settings like encryption only, CBC-MAC integrity-only, counter encryption only and counter encryption - CBC-MAC integrity with dynamic and static key support.

The paper is outlined as follows. Section 2 presents the background of various security protocols at different levels and their limitations, Section 3 formulates the implementation of flexible framework and its design goals, and introduces software AES and various modes of operation used by us and Section 4 describes the results and analysis of framework with software AES. Finally Section 5 concludes with a peek into the future work.

II. BACKGROUND

A. *Existing security protocols:*

TinySec [3] which is implemented at Link layer ensures low energy consumption but the algorithm is vulnerable to replay attacks. MiniSec [6] at Network layer ensures low energy consumption but operates only a fixed level of security which supports both encryption and authentication. Many applications require a combination of both confidential and non-confidential data which is not supported by MiniSec. However hardware AES security hosts ensures low power consumption and combinational levels of security. This feature is available for hardware that contains IEEE 802.15.4 compliant RF transceiver like CC2420/CC2430/CC2431 chips [6] and some of the motes that support this radio are TelosB, MicaZ and Mote2. Not all motes like IRIS [19] etc support hardware security. A software implementation of the AES could be a way satisfies these requirements widely. We have chosen AES for both efficiency and security reasons (see [8, 9]).

B. *Supported block ciphers:*

There are many block ciphers available namely Skipjack [16], RC5 [15], RC6 [11] and Rijndael [17]. Each cipher is chosen based on the need of applications security, memory and energy efficiency of the cipher [10].We have selected Rijndael in the configuration of 128/128/10 (keysize/blocksize/rounds) but still our library suite supports [192—256]/128/[12—14] configurations. As, in WSN power is more of concern, choosing 128 bit key is more appropriate.

C. *Modes of Block cipher:*

CTR encryption: This is counter mode encryption. To make compatible with hardware AES we provided this option [18].

CBC-MAC Authentication: This mode provides only authentication of the payload [7].

CCM Mode: This is Counter mode encryption and CBC-MAC authentication mode that features authenticated encryption [20].

OCB Mode: This Offset CodeBook, is a block-cipher mode of operation that features authenticated encryption for arbitrary length of data [7].

Only-Encryption/Only-Decryption: In this mode, simple encryption/decryption operations are performed without any mode settings.

III. THE ADAPTIVE FRAMEWORK FOR SECURE WSN APPLICATIONS AND IT'S DESIGN GOALS

We propose a flexible framework for programming TinyOS. This framework is divided into multiple components as shown in Figure 1 and each component contains group of protocols: routing protocols, Time-synchronization protocols, localization and security protocols. The application developer is required to wire a combination of these protocols for specific application using the framework. For example, an application can use TinyHop routing algorithm, Flooding Time Synchronization Protocol (FTSP) [9] with Hardware or Software security. In such a case this framework can be configured depending on the application's need and the level of security required. So, we grouped these protocols into multiple components by providing abstraction to the application developer for easier access.

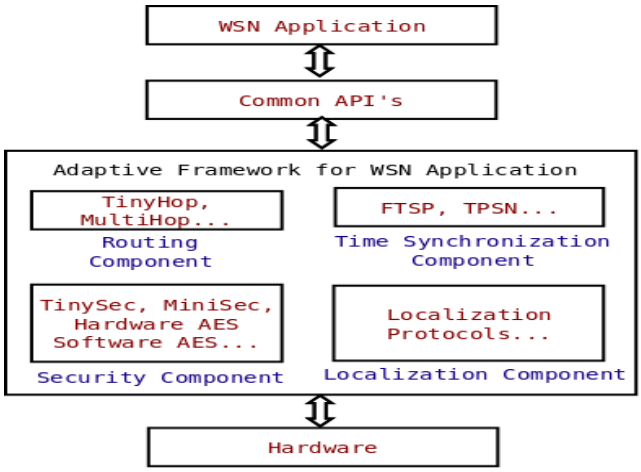


Fig. 1. Adaptive Framework

Since one of the objectives of this paper is in providing a unified security access API to application developers, we introduce a use case of the framework with the security component described. This security library lists available security algorithms, based on the version of TinyOS and the hardware platform chosen. There are two actors in the framework, the application developer and the system component developer. A system component developer

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

integrates security algorithms into the library suite in conjunction with the unified access API which we have standardized. The system component developer is required to create glue logic between the standardized API and the algorithm's internal functions. The application developer on the other hand is required to use the unified access API to wire his application's logic with the required security algorithm. Algorithms implementing various modes of block ciphers discussed earlier are integrable into the security suite. We have integrated our software AES with IEEE 802.15.4 specification and supported all modes for existing security algorithms is outlined in Table V. We have also integrated security algorithms like TinySec [3], MiniSec [6], and Hardware Security [5] in the suite. This framework is useful for security related experimentations in perspective of both application developer as well as System developer. Any new cipher can be easily plugged-in and plugged-out. The unified APIs provided for existing security algorithms removes the need of changing the existing application code. The framework takes care of mode settings based on chosen security layer and algorithm. This framework is simple in ease of use, flexible and adaptive.

TABLE I
WRAPPER APIS PROVIDED BY FRAMEWORK

```
command error t AFWAupdateKey(uint8_ t * key, set KEY);
event void AFWAupdateKeyDone(uint8_ t * key);
command error t AFWAsetTransmitMode( uint16_ t ctrl0,
                                     uint8_ t len );
command error t AFWAsetReceiveMode (  uint16_ t ctrl0,
                                     uint8_ t len );
command error t AFWAsend( uint16_ t addr,
                          message_ t *msg,  uint8_ t len  );
event void AFWAsendDone( message_ t *msg,
                         uint8_ t error);
command error t AFWAreset(uint8_ t Type);
command error t AFWAget(uint8_ t *key, get KEY);
```

Ease of use: - It provides complete abstraction to the application developer and introduces a GUI through which the application developer can select components of the lower layers. The framework will internally wire and create a template for application development. For example, if application developer uses IRIS mote and needs software AES with various mode settings, a project can be created for IRIS mote using WSN IDE [21] and then call commands for mode settings. Now the framework makes a setup and loads AES library.

Flexible: - Our framework is flexible, because it has feature of plug-in and plug-out facility i.e. any security algorithm can easily be integrated. Also, there are several schemes for key settings. In such a case dynamic key support is more robust than the pre-configured key. We used Java Cryptographic [23] functions with mouse movement and keyboard random-key generation. In this framework it supports both randomly generated dynamic key and static key setting.

Adaptive: - Level of security is application specific. So, it is a choice of application developer to choose type of security needed. The framework is adaptive so that it can switch between levels of security and provides corresponding security layer to the application based on selection of security algorithm. For example, applications may use TinySec or MiniSec, Hardware AES or Software AES etc.

A. Hardware Independent AES with IEEE 802.15.4 Specification

In this section, implementation of the Software AES security architecture with IEEE 802.15.4 specification which is provided in the framework security component is outlined. According to IEEE 802.15.4 specification it has eight different security suites [13]. The Table IV, gives five modes and in last two modes each have different variants based on the chosen MAC value.

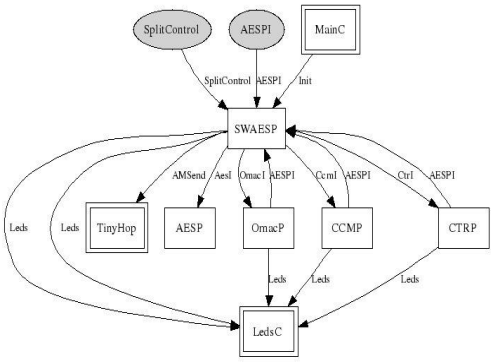


Fig. 2. Configuration file for Software AES

The above Figure 2 depicts the configuration for software AES. We have configured in such a way that n-number of modes can be integrate with AES module but only suitable mode with AES module is loaded during the compile time. This configuration makes developer easier to integrate AES module with any new mode.

We will see each AES security suite in-detail:

- **CTR:** This is counter mode encryption. It uses counter value and it consists of sender's address and 4-byte frame counter. We have not appended any flags and block counter to the data payload and hence this minimizes the power of mote. The block counter is the number of blocks splited into 16 byte blocks within the packet. This value can be calculated based on the size of the packet that is to be sent. The frame counter is maintained by the software AES and it is incremented for each packet automatically by software. The send API includes frame counter and encrypted payload into the data payload of the packet. Below is the code snippet for CTR mode encryption.

```
uint8_t *payload = call AMSend.getPayload(msg, len);
memcpy( smsg.data, payload, len );
call CtrI.AESctr encrypt((uint8_ t *)smsg.data, len,
                         sec txl, nonceValue );
memcpy(&fc1, &nonceValue[3], 4 );
smsg.fc = fc1;
memcpy(payload, &smsg, fcLen + len );
return call AMSend.send( addr, msg, fcLen + len );
```


http://sites.google.com/site/ijcsis/
ISSN 1947-5500

For example, if $h \parallel p_1,...,p_n$ is frame format then after CTR mode encryption, packet sent will be $h \parallel z_i \parallel c_i$, where, $c_i = p_i \oplus AES_k(x_i)$, h = *header of TinyHop*, z_i = *Frame counter*, c_i = *encrypted payload*, x_i = *nonce value or counter or initial vector*.

- **CBC_MAC‖SEC_M [0 − 7]:** This mode provides only authentication. And the size of MAC value can vary between 4, 6, 8, 10, 12, 14, 16 byte. Here CBC_MAC‖SEC_M[0− 7] is a macro setting that tells library to load AES with CBC_MAC mode. In the below code snippet *truncateTag()* command will truncates MAC in the range {4,6,8,10,12,14,16} byte based on the SEC_M[0− 7] macro setting i.e. if the macro is set to one among the following SEC_M1, SEC_M2 ... SEC_M7 then each value maps to the range {4,6,8,10,12,14,16} byte while SEC_M0 is reserved for future use. By default in CBC-MAC mode, length field is authenticated by software.

  In this it supports two protecting modes:

  (a) It protects header of TinyHop routing algorithm as well as the data payload. Suppose if $h \| p_1,...,p_n$ is frame format then after CBC_MAC authentication, packet that is sent will be

  $$h \| p_1,...,p_n \parallel Trunc_{SEC\_M[0-7]}\{auth(h \| p_1,...,p_n)\}.$$

  (b) Protects only data payload. Suppose if $h \| p_1,...,p_n$ is frame format then after CBC_MAC authentication, packet that is sent will be

  $$h \| p_1,...,p_n \parallel Trunc_{SEC\_M[0-7]}\{auth(p_1,...,p_n)\}.$$

  Where, h = *TinyHop header*, p_i= *plain text*.

```
uint8_t *payload = call AMSend.getPayload(msg, len);
memcpy( smsg.data, payload, len );
call CtrI.AESctr encrypt((uint8_ t *)smsg.data, len,
                         sec txl, nonceValue );
memcpy(&fc1, &nonceValue[3], 4 );
smsg.fc = fc1;
memcpy(payload, &smsg, fcLen + len );
return call AMSend.send( addr, msg, fcLen + len );
```

- **CCM‖SEC_M [0− 7]:** This is AES counter mode encryption and CBC-MAC authentication. Here CBC_MAC‖SEC_M[0 − 7] is a macro setting that tells library to load AES with CCM mode and functionality of SEC_M[0− 7] is same as previous mode setting. In this mode first it authenticates header of TinyHop routing algorithm and data payload using CBC-MAC and then encrypts both data payload and MAC using AES-CTR mode. Below is the code snippet for CCM mode implementation.

```
uint8_t *payload = call AMSend.getPayload(msg,
                                          len);
memcpy(smsg.data, payload, len );
call CcmI.AESccm nonce(nonceValue);
call CcmI.AESccm auth((uint8_ t *)smsg.data, len,
                      sec txl,KeySizeB );
call CcmI.truncateTag((uint8_ t *)smsg.data,len,
                      appLen);
call CcmI.AESccm encrypt((uint8_ t *)smsg.data, len,
                         sec txl,nonceValue );
memcpy( &fc1, &nonceValue[3], 4 );
smsg.fc = fc1;
memcpy( payload, &smsg, fcLen + len + appLen );
return call AMSend.send( addr, msg, fcLen + len +
                         appLen );
```

For example, If $h \| p_1, \ldots, p_n$ is frame format then after CCM mode, packet sent will be

$$h \parallel z_i \parallel c_1, \ldots, c_n \parallel ENC(MAC)$$

where, $MAC = Trunc_{SEC\_M}[0-7]\{auth(h \| p_1, \ldots, p_n)\}$, $ENC(MAC) = MAC \oplus AES_k(x_i)$, $c_i = p_i \oplus AES_k(x_i)$, z_i = *Frame Counter*.

- **AES_ENC:** This is simplest mode. It provides simple AES encryption operation without any mode settings. Below is the code snippet that takes key size and pointer to input array as an argument and produces encrypted output in the same input array.

```
call AesI.startAES((uint8_ t)KeySize, (uint8_ t *)inPut);
```

- **AES_DEC:** This is simplest mode. It provides simple AES decryption operation without any mode settings. Below is the code snippet that takes key size and pointer to input array as an argument and produces decrypted output in the same input array.

```
call AesI.startAES((uint8_ t)KeySize, (uint8_t *)inPut);
```

Figure 3, depicts the complete flow that takes changes in the length field of the payload while sending a packet.

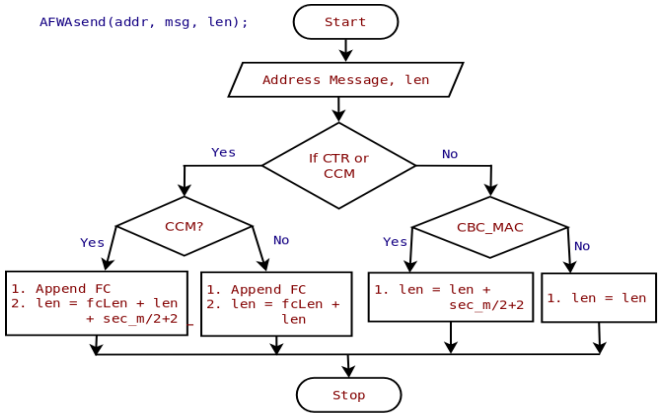


Fig. 3. AES Security suite

FC = *frame counter*, fcLen = *frame counter length*
len = *length of the payload*
sec_m = *Number of bytes in authentication field for CBC-MAC, encoded as (M-2)/2.*

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

B. Common API's for Security algorithms

Since, each security algorithm has its own parameter settings; the developer has to know which API is to be used. We propose Security interface that contains common wrapper API's to set and get parameters for various security algorithms. Before choosing security algorithm, application developer will choose version of tinyos and for this we created framework [21] that lays above wrapper API's. Some of the common API's that we provided are:

*a) command error t AFWAupdateKey(uint8_t * key, uint8_t setKEY);*

*event void AFWAupdateKeyDone(uint8_t * key);*

AFWAupdateKey() updates old key value to new key value where KEY parameter defines which key is to be updated in ACL entry. *uint8_t *key* is a pointer to an array containing 8-bit unsigned integers. KEY is 8-bit unsigned integer that takes macro value for ACL entries i.e. KEY0 or KEY1 or so on.

b) command error t AFWAsetTransmitMode(uint16_t ctrl0, uint8_t len);
command error t AFWAsetReceiveMode (uint16_t ctrl0, uint8_t len);

AFWAsetTransmitMode() / AFWAsetReceiveMode() sets transmission and receiver mode for any security algorithm. Parameter *ctrl0* value can be a combination of macros given in Table IV and parameter *len* has value zero if the total payload value is to be encrypted/decrypted otherwise *len* value specifies number of bytes to be encrypted or decrypted or number of bytes not to be encrypted or decrypted based on the context and mode setting of *ctrl0*.

*c) command error t AFWAsend(uint16_t addr, message_t *msg, uint8_t len);*
*event void AFWAsendDone(message_t *msg, uint8_t error);*

AFWAsend( ) command is similar to AMSend() command. This command will set correct length of the transmitted message when Hardware/Software AES security is used and this command does nothing for other security algorithms.

d) command error t AFWAreset(uint8_t Type);

This command is used to reset MAC or Encryption initialization vectors of security algorithm.

*e) command error t AFWAget(uint8_t *key, KEY);*

This command is used to get key value from ACL entry. KEY is 8-bit unsigned integer that takes macro value for ACL entries i.e. KEY0 or KEY1 or so on and the final result is fetched to key.

IV. Results and Analysis

We have tested our proposed framework and software Advanced Encryption Standard (AES) with various mode settings using two different motes i.e. IRIS and MicaZ. We have taken application that sends encrypted packets using TinyHop routing algorithm to the Base Station. And then Base Station decrypts the received packets and forwards the packet to serial forwarder where the user can view the original packet. We sniffed the packets using sniffer to check whether packet is encrypted or not in various mode settings. Tables II results are obtained after installing software Advanced Encryption Algorithm using TinyOS-2.1.0 without TinyHop routing algorithm and without our framework into MicaZ and IRIS motes.

TABLE II
Memory Utilization of Software AES using TinyOS-2.1.0 Framework Without Tinyhop Routing Algorithm

Mote	ROM occupied in bytes (percentage)	RAM occupied in bytes (percentage)	Name of Cipher and its configuration
MicaZ	22256(17.38%)	2196(54.9%)	AES128/128/10, CCM mode with 16- byte
	22258(17.38%)	2196(54.9%)	AES 128/128/10 CBC-MAC with 16-byte MAC
	22254(17.38%)	2196(54.9%)	AES 128/128/10 CTR mode
	22242(17.37%)	2196(54.9%)	AES 128/128/10 only AES Encryption/ Decryption
IRIS	21308(16.64%)	2404(30.05%)	AES 128/128/10 CCM mode with 16- byte
	21306(16.64%)	2404(30.05%)	AES 128/128/10 CBC-MAC with 16-byte MAC
	21304(16.64%)	2404(30.05%)	AES 128/128/10 CTR mode
	21292(16.63%)	2404(30.05%)	AES 128/128/10 Only AES Encryption/ decryption

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Table III results are obtained after installing software Advanced Encryption Algorithm using TinyOS-2.1.0 with TinyHop routing algorithm and with our framework into MicaZ and IRIS mote. The TinyHop routing algorithm occupies more memory as to manage routing information. Also we have removed distinction between key and frame counter [13].

TABLE III
MEMORY UTILIZATION OF SOFTWARE AES USING TINYOS-2.1.0 FRAMEWORK WITH TINYHOP ROUTING ALGORITHM

Mote	ROM occupied in bytes (percentage)	RAM occupied in bytes (percentage)	Name of Cipher and its configuration
MicaZ	29848(23.31%)	3739(93.4%)	AES 128/128/10 CCM mode with 16-byte
	29850(23.32%)	3663(91.5%)	AES 128/128/10 CBC-MAC with 16-byte
	29774(23.26%)	3435(85.87%)	AES 128/128/10, CTR mode
	29676(23.18%)	3359(83.97%)	AES 128/128/10, Only AES Encryption/ Decryption
IRIS	28758(22.46%)	3923(49.03%)	AES 128/128/10, CCM mode with 16-byte MAC
	28760(22.46%)	3831(47.88%)	AES 128/128/10, CBC-MAC with 16-byte MAC
	28682(22.40%)	3555(44.43%)	AES 128/128/10, CTR mode
	28650(22.38%)	3359(41.98%)	AES 128/128/10, only AES Encryption/ Decryption

V. CONCLUSION AND FUTURE WORK

The framework implemented is unique for WSN applications. It has general purpose security library suite composed of existing security algorithms for popular sensor node hardware platforms. The framework also supports existing versions of TinyOS providing the application developer an abstraction to low level implementation of the desired security algorithm chosen. We also Implemented and integrated software AES and hardware AES driver for CC2420/CC2430/CC2431 radio chip [5] with same mode settings like encryption-only, CBC-MAC integrity-only, counter encryption only and counter encryption – CBC-MAC integrity with dynamic and static key support. We have tested the code using with and without TinyHop routing algorithm.

The work presented in this paper can be applied to various mote platforms. Future extensions are also possible to integrate various security modes with simple plug-in and plug-out configuration file. Currently the framework is in-built with WSN IDE [21] that creates templates for application and then application developer has to wire the modules manually. In the next design we extend this framework and security modules wiring with drag and drop options.

REFERENCES

[1] K. Piotrowski, P. Langendoerfer, S. Peter,"How public key cryptography influences wireless sensor node lifetime," *Proc. of fourth ACM workshop on Security of adhoc and sensor networks*(SASN 06), pages 169-176,2006. ACM..

[2] A. Wander, N. Gura, H. Eberle, V. Gupta, S. C. Shantz,"Energy analysis of public-key cryptography for wireless sensor networks," *Proc. of Third IEEE International Conference on Pervasive Computing and Communications* (PerCom 05), pages 324-328, March 2005.

[3] Chris Karlof, Naveen Sastry, David Wagner, "TinySec: Link Layer Encryption for Tiny Devices", *ACM Conference on Embedded Networked Sensor Systems*, 2004.

[4] Tieyan Li, Hongjun Wu, Xinkai Wang, Feng Bao,"SenSec Design, I2R Sensor Network Flagship Project"; Technical Report TRv1.0.

[5] 2.4 GHz IEEE 802.15.4/ZigBee-ready RF Transceiver, Chipcon Products from Texas Instruments. http://www.ti.com.

[6] Mark Luk, GhitaMezzour, Adrian Perringm, Virgil Gligor,"MiniSec: A Secure Sensor Network Communication Architeture", *ACM International Conference on Information Processing in Sensro Networks*, April 2007.

[7] Phillip Rogaway, Mihir Bellare, John Black, "OCB: A block-cipher mode of operation for efficient authenticated encryption", *ACM Transactions on Information and System Security* (TISSEC), Volume 6, Issue 3, pp.365-403, August 2003.

[8] Devesh Jinwala, Dhiren Patel, K S Dasgupta,"Optimizing the Block Cipher Modes of Operations Overhead at the Link Layer Security Framework in the Wireless Sensor Networks," *Proceedings of the 4th International Conference on Information Systems Security*, LNCS, pp.258-272, Springer Berlin/Heidelberg, 2008.

[9] Miklos Maroti, Gyula Simon, Branislav Kusy, and Akos Ledeczi,"The flooding time synchronization protocol," in *Proceedings of the 2nd international conference on Embedded networked sensor systems*, Baltimore, MD, USA, Nov. 2004, pp. 3949.

[10] Law, Y.W., Doumen, J., Hartel, P, "Survey and benchmark block ciphers for wireless sensor networks"; *ACM Transactions on Sensor Networks*, 2006.

[11] RC6 cipher - http://people.csail.mit.edu/rivest/Rc6.pdf

[12] Mingbo Xiao, Xudong Wang, Guangsong Yang,"Cross-Layer Design for the Security of Wireless Sensor Networks", *Proceedings of the 6th World Congress on Intelligent Control and Automation*, Jnue 21-23,2006 Dalian, China,pp(104-108).

[13] Naveen Sastry and David Wagner, "Security Considerations for IEEE 802.15.4 Networks". *ACM Workshop on Wireless Security WiSe* 2004, October 2004.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

[14] Kalpana Sharma, M.K. Ghose, Kuldeep, "Complete Security Framework for Wireless Sensor Networks", (IJCSIS) *International Journal of Computer Science and Information Security*, Vol. 3, No. 1, 2009.

[15] Rivest R; "The RC5 Encryption Algorithm", *Proceedings of the Second International Workshop on Fast Software Encryption*, 1994.

[16] Skipjack - a representative of a family of encryption algorithms as part of the NSA suite of algorithms; http://csrc.nist.gov/groups/STM/cavp/documents/skipjack/skipjack.pdf

[17] J.Daemen, V.Rijmen, "AES Proposal: Rijndael", http://www.esat.kuleuven.ac.be/rijmen/rijndael/rijndaeldocV2.zip.

[18] Helger Lipmaa, Phillip Rogaway and David Wagner. Comments to NIST Concerning AES-modes of Operations: CTR-mode Encryption. In Symmetric Key Block Cipher Modes of Operation Workshop, Baltimore, Maryland, USA, October 20, 2000.

[19] IRIS motes - http://www.xbow.com/Products/wproductsoverview.aspx.

[20] Whiting, D., Housley, R. and N. Ferguson, "AES Encryption Authentication Using CTR Mode CBC-MAC," *IEEE P802.11 doc* 02/001r2, May 2002.

[21] http://www.ubicomp.in/afwa/

[22] http://www.tinyos.net/

[23] http://java.sun.com/j2se/1.4.2/docs/guide/security/CryptoSpec.html

TABLE IV
MODE SETTINGS FOR VARIOUS SECURITY ALGORITHMS

Algorithm	ctrl0	len
AES: Stand Alone	AES_STANDALONE ‖KEY0	0
	AES_STANDALONE ‖KEY1	0
AES: In-line	AES_INLINE ‖[T‖R]XKEY [0|1] ‖CBC_MAC‖SEC_M[0 − 7]	X
	AES_INLINE ‖[T‖R]XKEY[0|1] ‖CT R	X
	AES_INLINE ‖[T‖R]XKEY [0|1] ‖CCM	X
	AES_INLINE ‖[[T‖R]XKEY[0|1] ‖CBC_MAC ‖SEC_M[0 − 7]	X
TinySec	TINYSEC_AUTH_ONLY	0
	TINYSEC_ENCRYPT_AND_AUTH	0
	TINYSEC_DISABLED	0
	TINYSEC_RECEIVE_ AUTHENTICATED	0
	TINYSEC_RECEIVE_CRC	0
	TINYSEC_RECEIVE_ANY	0
MiniSec	MINISECU	0

TABLE V
MODE SETTINGS FOR VARIOUS SECURITY ALGORITHMS

Algorithm	Macro	Description
AES	AES_ENC	only AES encryption
	AES_DEC	only AES decryption
	CTR	AES Counter-Mode Encryption This mode is not secure, to make compatible with hardware AES we provided this option.
	CBC_MAC ‖SEC_M[0-7]	4, 6, 8,10,12,14, 16 byte -MAC. AES CBC-MAC, it provides only Authentication.
	CCM ‖SEC_M[0-7]	4,6,8,10,12,14, 16-byte MAC. AES Counter mode encryption and CBC-MAC authentication

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Optimizing the Application-Layer DDoS Attacks for Networks

P.Niranjan Reddy
Head, Dept. of CSE
KITS, Warangal
A.P. , INDIA.
npolala@yahoo.co.in

K.Praveen Kumar
Lecturer, Dept, of CSE
KITS, Warangal
A.P. , INDIA.
praveen_kumar35@yahoo.co.in

M .Preethi
Lecturer, Dept, of CSE
KITS, Warangal
A.P. INDIA.
preethi_0290@yahoo.co.in

***Abstract* – The main aim of the proposed framework is to implement the Application-Layer DDoS Attacks Optimizing for Popular Websites that employing legitimate HTTP requests to flood out victim resources and to implement an effective method to identify whether the surge in traffic is caused by App-DDoS attackers or by normal Web surfers.**

Keywords: Terms – Application-layer, distributed denial of service (DDoS), popular website.

I. INTRODUCTION

Distributed Denial of Service (DDoS) attack is an attempt to make a computer resource unavailable to its intended users. This attack has caused severe damage to servers and will cause even greater intimidation to the development of new Internet services. Traditionally, DDoS attacks are carried out at the network layer, such as ICMP flooding, SYN flooding, and UDP flooding, which are called Net-DDoS attacks. The intent of these attacks is to consume the network bandwidth and deny service to legitimate users of the victim systems. Among these floodings another attack is Botnet[21] which is a network of compromised hosts or *bots*, under the control of a human attacker known as the botmaster. Botnets are used to perform malicious actions, such as launching DDoS attacks, sending spam or phishing emails and so on. Thus, botnets have emerged as a threat to internet community. Peer to Peer (P2P) is a relatively new architecture of botnets. These botnets are distributed, and small. So, they are difficult to locate and destroy.

Since many studies have noticed this type of attacks and have proposed different schemes (e.g., network measure or anomaly detection) to protect the network and equipment from bandwidth attacks, it is not as easy as in the past for attackers to launch the DDoS attacks based on network layer. To implement DDoS, a worm like program is created to simulate self-propagation onto many hosts on a network.

When the simple Net-DDoS attacks fail, attackers shift their offensive strategies to application-layer attacks and establish a more sophisticated type of DDoS attacks.

To overreach detection, the attackers attacking the victim web servers by HTTP GET requests (e.g., HTTP flooding) and pulling large image files from the victim server in overwhelming numbers. In another instance, attackers run a massive number of queries through the victim's search engine or database query to bring the server down [4]. Such attacks called as application-layer DDoS (App-DDoS) attacks. The MyDoom worm [23] and the CyberSlam [3] are all instances of this type attack.

On the web, "flash crowd"[6],[7] refers to the situation when a very large number of users simultaneously access a popular website[13], which produces a surge in traffic[8] to the Website and might cause the site to be virtually unreachable. Because burst traffic and high volume are the common characteristics of App-DDoS attacks and flash crowds, it is not easy for current techniques to distinguish them merely by statistical characteristics of traffic.

II. RELATED WORK

The researchers made an attempt to detect DDoS attacks from three different layers: IP layer, TCP layer, and application layer. From all of these views, researchers are looking into various approaches to differentiate normal traffic from the attack one.

Maximum DDoS-related research has concentrated on the IP layer. These techniques attempt to detect attacks by analyzing specific features, e.g., arrival rate or header information. For example, Cabrera *et al.* [9] used the management information base (MIB) data which include parameters that indicate

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

different packet and routing statistics from routers to achieve the early detection. Yuan *et al.* [14] used the cross-correlation analysis to capture the traffic patterns and then to decide where and when a DDoS attack possibly arises. Mirkovic *et al.* [15] monitored the asymmetry of two-way packet rates and to identify attacks in edge routers. Other statistical approach for detection of DDoS attacks includes IP addresses [16] and time-to-live (TTL) values [17].

One of the most important research area is TCP layer for detecting DDoS attack. For example, authors [9] mapped ICMP, UDP, and TCP packet statistical abnormalities to specific DDoS attacks based on MIB. Wang *et al.* [18] used the TCP SYN/FIN packets for detecting SYN flooding attacks. In [18], DDoS attacks were discovered by examining the TCP packet header against the welldefined rules and conditions and differentiated the difference between normal and abnormal traffic. Noh *et al.* [19] attempted to detect attacks by computing the ratio of TCP flags (including FIN, SYN, RST, PSH, ACK, and URG) to TCP packets received at a Web server.

Ranjan *et al.* [11] used statistical methods to detect characteristics of HTTP sessions and employed rate-limiting as the primary defense mechanism. Yen *et al.* [12] defended the application DDoS attacks with constraint random request attacks by the statistical methods. Other researchers combated the App-DDoS attacks by "puzzle," see, e.g., [20]. Jung *et al.*'s work [7] he used two properties to distinguish the DoS and normal flash crowd: 1) a DoS event is due to an increase in the request rates for a small group of clients while flash crowds are due to increase in the number of clients and 2) DoS clients originate from new client clusters as compared to flash crowd clients which originate from clusters that had been seen before the flash event.

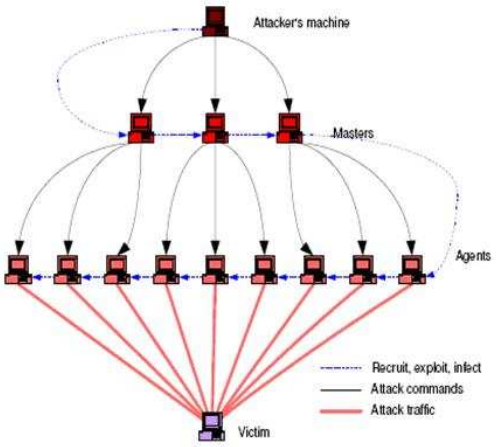


Fig 1. How the attacker can perform attacks on App-layer.

III. App-DDoS ATTACKS

In our opinion, the DDoS attack detection approaches in different scenario can be clustered as:

- Net-DDoS attacks versus stable background traffic.
- Net-DDoS attacks versus flash crowd.
- App-DDoS attacks versus stable background traffic.
- App-DDoS attack versus flash crowd.

The first two scenarios have been well studied and can be dealt with by most existing DDoS detection schemes while the other two groups are quite different from the previous ones.

This is a simple comparison between the existing system and proposed system.

Existing System	**Proposed System**
Consume the network bandwidth and deny	Bandwidth is effectively used
Service to legitimate users.	Service to all users if and only if the resource is available.
Abnormalities are identified and denied	Identifying abnormalities and serve them in different priority queues.
Large amount of data is required to train.	Identifies abnormalities with small amount of training data
Only positive data's are used to train	More accurate identification
Identifying abnormal traffic and filter the network	Identifying most abnormal traffic and filter when the network is heavily loaded.

IV. DETECTION PRINCIPLE

We can cluster the Web surfers and evaluate their contributions to the anomalies in the aggregate Web traffic. Here the DDoS attack is caused only by the authenticated users of the Website. Then, different priorities are given to the clusters according to their abnormalities and serve them in different priority queues. The most abnormal traffic may be filtered when the network is heavy loaded. Here the priority level of the cluster is given based on the access time only. The

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

different modules in the implementation are given below in the next section.

V. MODULES

Web Server Module

- Login
- Registration
- Database Design
- Application Design

Attacker Module

- Normal User
- Abnormal User

Flash crowd dismisser

- Data preparation
- Training
- Monitoring

A. Web Server Module

Web servers are computers on the internet that host website serving pages to viewers upon request. This service is referred to as web hosting. Every web server has a unique address so that other computers connected to the internet know where to find it on the vast network. When your request reaches its destination, the web server that hosts website sends the page in HTML code to your ipaddress [5]. This return communiqué travels back through the network. Your computer receives the code and your browser interprets the HTML code then displays the page for you in graphic form.

B. Login

Login module is general for all kinds of Web application to authenticate and authorize the user's access to the site. To make valid users only can access the site, preventing the unauthorized access.

C. Registration:

This module is also common to all the web application. Making the users to access the site based upon the registration. It may be free or cost. In order to authenticate and authorize a user, registration is must.

D. Application Design:

An application which suits for our project is designed using the HTML code and the relevant technologies.

E. Database Design:

Once the application has designed then Database has to be designed. Here creation of the tables related to our project is created. Number of tables needed for the application has to be decided and the tables are created for that.

F. Attacker Module:

This module consists of webpage through which Attackers attack the victim Web servers by HTTP GET requests (e.g., HTTP Flooding) and pulling large image files from the victim server in overwhelming numbers. In another instance, attackers run a massive number of queries through the victim's search engine or database query to bring the server down. Very large number of attackers simultaneously accesses a popular Website, which produces a surge in traffic to the website and might cause the site to be virtually unreachable.

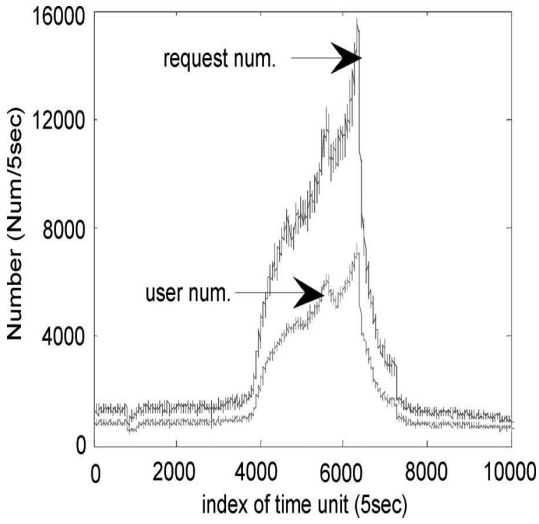


Fig 2. Simple network Attack path

Normal User: The user login in and acts as a normal user there is no abnormality in his behaviors.

Abnormal User: The user login in and acts as the abnormal user, the behaviors of the users are found to be abnormal (.e.g., attacker who is causing the DDoS attack over the target site).

Flash crowd dismisser: This model is first trained by the stable and low-volume web workload whose normality can be ensured by most existing anomaly detection systems, and then it is used to monitor the following web workload for a period of 10 min. When the period is past, the model will be updated by the new collected web workload whose normality is ensured by its entropy fitting to the model. Then, the model is used in anomaly detection for the next cycle. If some abnormalities hiding in the incoming web traffic are found, the "defense" system will be implemented.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

VI. ARCHITECTURE

The process is divided into three phases:

1. Data preparation.
2. Training
3. Monitoring

Data preparation: The main purpose of data preparation is to compute the AM by the logs of the web server. Various user data are collected while accessing the sites.

Training: Train the collected data for the abnormalities. Check the user behaviour with the predefined threshold. If the user exists the threshold are named as the abnormal users (.eg., attacker). Likewise all the user data are trained and found out the abnormality.

Monitoring: In the Monitoring phase, checks for the resource availability. If the user found to be attacker then the resource is available means allows that user to access the sites (Simply allow the attacker also if and only if the resource is available). If the resource is not available means, temporarily deny that user to access the site.

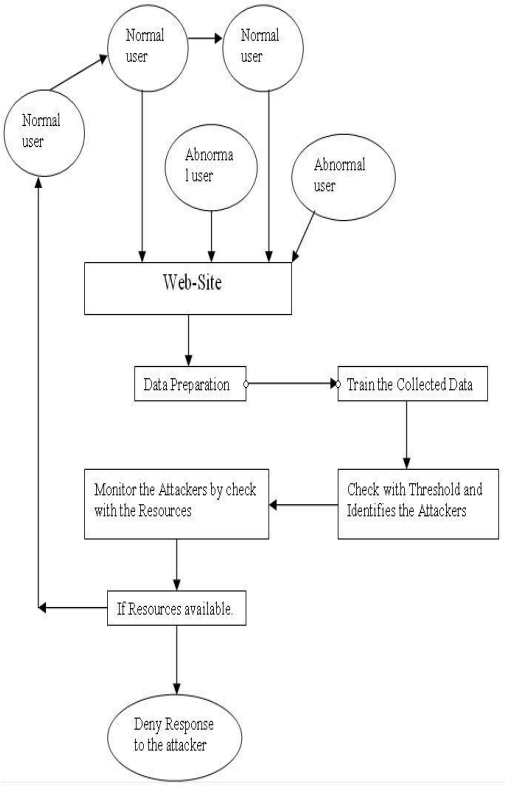


Figure 3. Proposed Architecture

VII. APPLICATION

Web servers application DoS attacks allow for efficient DoS with only little resources at hand, and thus pose a **Serious threat to organization**.

- Hide speed internet.
- Mobility tracking in wireless networks.

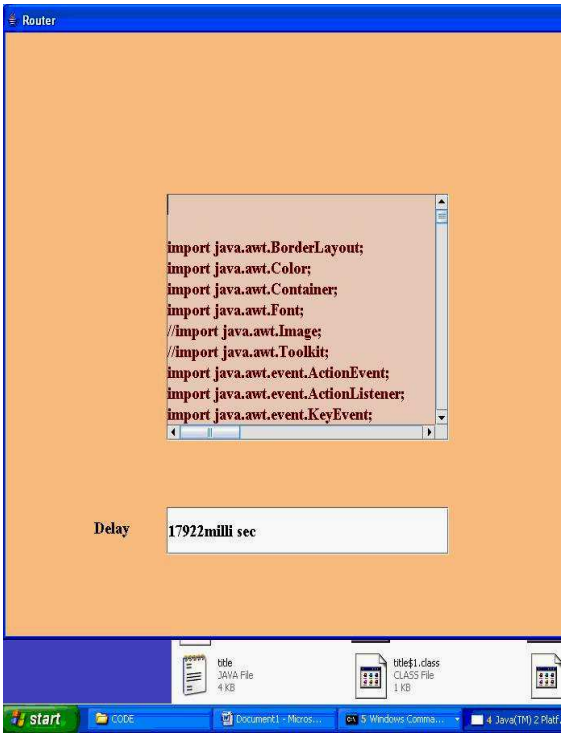


Figure 4. Time delay while transferring the file with out attack.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

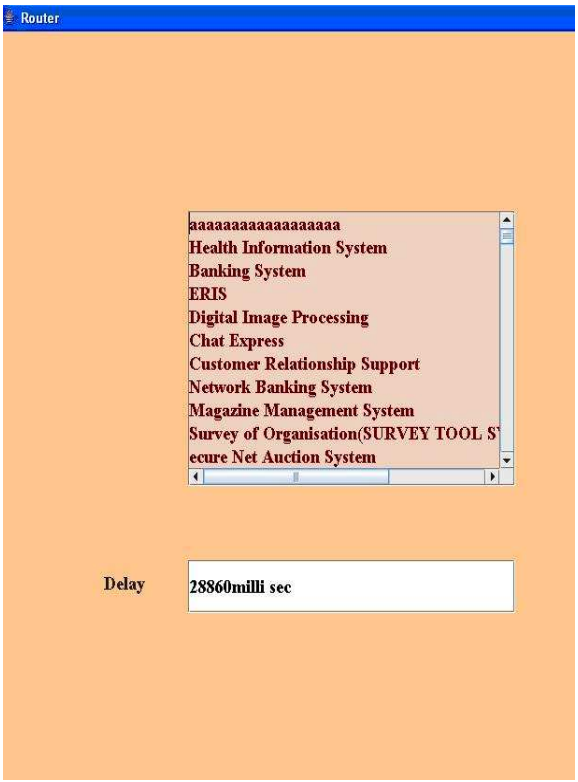


Figure 5. Time delay while transferring the file with attack.

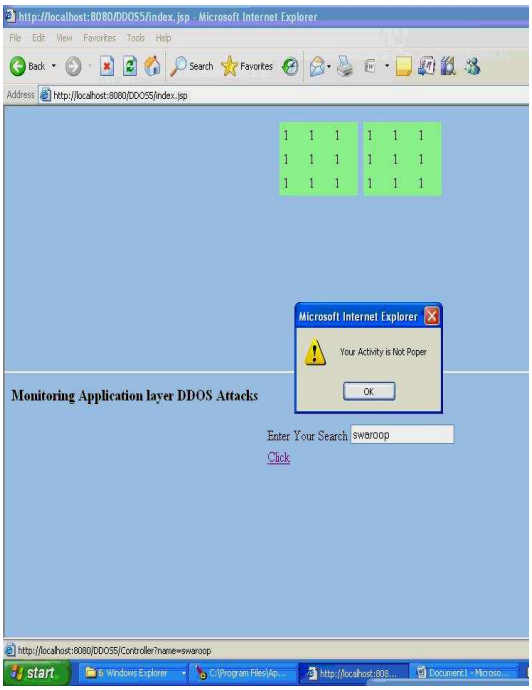


Figure 6. Session closing when the Attacking is found.

VIII. CONCLUSION

Creating defenses for attacks requires monitoring dynamic network activities in order to obtain timely and signification information. While most current effort focuses on detecting Net-DDoS attacks with stable background traffic, we proposed detection architecture in this paper aiming at monitoring web traffic in order to reveal dynamic shifts in normal burst traffic, which might signal onset of App-DDoS attacks during the flash crowd event. Our method reveals early attacks merely depending on the document popularity obtained from the server log.

REFERENCES

[1]. IEEE/ACM Transaction on Networking, Vol. 17, No. 1, February, 0209.

[2]. Http://www.linuxsecurity.com/resource_files/intrusion_detection/ddos-whitepaper.html.

[3]. http://en.wikipedia.org/wiki/Denial-of-serviceattack.

[4]. K. Poulsen, "FBI Busts Alleged DDoS Mafia," 2004.[Online].Available: http://www.securityfocus.com/news/9411

[5]. T. Peng and K. R. M. C. Leckie, "Protection from distributed denial of service attacks using history-based IP filtering," in *Proc. IEEE Int. Conf. Commun.*, May 2003, vol. 1, pp. 482–486.

[6]. I. Ari, B. Hong, E. L. Miller, S. A. Brandt, and D. D. E. Long, "Modeling, Analysis and Simulation of Flash Crowds on the Internet," Storage Systems Research Center Jack Baskin School of Engineering University of California, Santa Cruz Santa Cruz, CA, Tech. Rep. UCSC-CRL-03-15, Feb. 28, 2004 [Online]. Available: http://ssrc.cse.ucsc.edu/, 95064.

[7]. J. Jung, B. Krishnamurthy, and M. Rabinovich, "Flash crowds and denial of service attacks: Characterization and implications for CDNs and web sites," in *Proc. 11th IEEE Int. World Wide Web Conf.*, May 2002, pp. 252–262.

[8]. W. Leland, M. Taqqu, W. Willinger, and D. Wilson, "On the selfsimilar nature of ethernet traffic (extended version)," *IEEE/ACM Trans. Networking*, vol. 2, no. 1, pp. 1–15, Feb. 1994.

[9]. J. B. D. Cabrera, L. Lewis, X. Qin, W. Lee, R. K. Prasanth, B. Ravichandran, and R. K. Mehra, "Proactive detection of distributed denial of service attacks using MIB traffic variables a feasibility

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

study," in *Proc. IEEE/IFIP Int. Symp. Integr. Netw. Manag.*, May 2001, pp. 609–622.

[10]. S. Noh, C. Lee, K. Choi, and G. Jung, "Detecting Distributed Denial of Service (DDoS) attacks through inductive learning," *Lecture Notes in Computer Science*, vol. 2690, pp. 286–295, 2003.

[11]. S. Ranjan, R. Swaminathan, M. Uysal, and E. Knightly, "DDoS-resilient scheduling to counter application layer attacks under imperfect detection," in *Proc. IEEE INFOCOM*, Apr. 2006 [Online].Available:http://www-ece.rice.edu/networks /papers/ dos-sched.pdf

[12]. W. Yen and M.-F. Lee, "Defending application DDoS with constraint random request attacks," in *Proc. Asia-Pacific Conf. Commun.*, Perth, Western Australia, Oct. 3–5, 2005, pp. 620–624.

[13]. C. Roadknight, l. Marshall, and D. Vearer, "File popularity characterisation," *ACMSIGMETRICS Performance Eval. Rev.*, vol. 23, no. 4, pp. 45–50, Mar. 2000.

[14]. J. Yuan and K. Mills, "Monitoring the macroscopic effect of DDoS flooding attacks," *IEEE Trans. Dependable and Secure Computing*, vol.2, no. 4, pp. 324–335, Oct.-Dec. 2005.

[15]. J. Mirkovic, G. Prier, and P. Reiher, "Attacking DDoS at the source," in *Proc. Int. Conf. Network Protocols*, 2002, pp. 312–321.

[16]. T. Peng and K. R. M. C. Leckie, "Protection from distributed denial of service attacks using history-based IP filtering," in *Proc. IEEE Int.Conf. Commun.*, May 2003, vol. 1, pp. 482–486.

[17]. B. Xiao, W. Chen, Y. He, and E. H.-M. Sha, "An active detecting method against SYN flooding attack," in *Proc. 11th Int. Conf. Parallel Distrib. Syst.*, Jul. 20–22, 2005, vol. 1, pp. 709–715.

[18]. H.Wang, D. Zhang, and K. G. Shin, "Detecting SYN flooding attacks," in *Proc. IEEE INFOCOM*, 2002, vol. 3, pp. 1530–1539.

[19]. S. Noh, C. Lee, K. Choi, and G. Jung, "Detecting Distributed Denial of Service (DDoS) attacks through inductive learning," *Lecture Notes in Computer Science*, vol. 2690, pp. 286–295, 2003.

[20]. S. Kandula, D. Katabi, M. Jacob, and A. W. Berger, "Botz-4-Sale: Surviving Organized DDoS Attacks that Mimic Flash Crowds,"MIT, Tech.Rep. TR-969, 2004 [Online]. Available: http://www.usenix.org/events/nsdi05/tech/ kandula/ kandula.pdf

[21]. James Binkley and Suresh Singh. An algorithm for anomaly-based botnet detection. In *Proceedings of Steps to Reducing Unwanted Traffic on the Internet Workshop (SRUTI '06)*, 2006.

[22]. Zhichun Li, Anup Goyal, and Yan Chen. Honeynet-based Botnet Scan Traffic Analysis Northwestern University, Evanston, IL 60208 {lizcag o210,ychen}@cs.northwestern.edu.

[23]. "Incident Note IN-2004-01 W32/Novarg. A Virus," CERT, 2004. [Online]. Available: http://www.cert.org/incident_notes/ IN-2004-01.html

AUTHORS PROFILE

P.NIRANJAN REDDY received the B.E Computer Science from Nagpur University in 1992 and M.Tech (Computer Science and Engineering) from NIT, Warangal in the year 2001.He worked as a Lecturer and Assistant Professor in the department of CSE of KITS, Warangal, Since 1996. He is doing a part-time research in Kakatiya University, Warangal since 2007. He authored two text books, Theory of computation and Computer Graphics in the field of Computer Science. He published 3 papers inInternational Journals and 6 papers in International Conferences.

K.PRAVEEN KUMAR has been working as a lecturer in Dept. of CSE, KITS,Warangal in Andhra Pradesh,INDIA for the last 2 years. He has completed his B.tech and M.tech from KITS warangal. He has published a research paper at a National level Conference.

M.PREETHI has been working as a lecturer in Dept of CSE in KITS, Warangal in Andhra Pradesh, INDIA for the last 3years. She took her M.Tech degree from KITS,Warangal.

Survey – New Routing Technique for Grid Computing

R.RAMESHKUMAR,
Research Schlolar,
J.N.T.University,
Kukatpally,
Hyderabad.
rramesh1968@gmail.com

Dr. A.DAMODARAM
Director/ U.G.C Academic Staff College,,
J.N.T.University,
Kukatpally,
Hyderabad.
adamodaram@jntuap.ac.in

Abstract- Trust plays an indispensable role in grid computing. Trust-management systems provide applications with a standard interface for getting answers to such questions and provide users with a standard language for writing the policies and credentials that control what is allowed and what isn't. Using a trust-management system for controlling security-critical services frees the application developer from a number of often difficult design and implementation issues and allows users to take advantage of a flexible, standard, application-independent language for specifying policy. In this paper, we develop trust management architecture for trust enhanced Grid security incorporating a novel trust model which is capable of capturing various types of trust relationships that exist in a Grid system and providing mechanisms for trust evaluation, recommendations and update for trust decisions. The outcomes of the trust decisions can then be employed by the Grid security system to formulate trust enhanced security solutions. Here we put forth ant algorithm for implementation. The ant colony algorithm is an algorithm for finding optimal paths that is based on the behavior of ants searching for food. At first, the ants wander randomly. When an ant finds a source of food, it walks back to the colony leaving "markers" (pheromones) that show the path has food. When other ants come across the markers, they are likely to follow the path with a certain probability. If they do, they then populate the path with their own markers as they bring the food back. As more ants find the path, it gets stronger until there are a couple streams of ants traveling to various food sources near the colony.

Key words - Grid computing, security, trust, Ant Colony, Service Request.

I. INTRODUCTION

Routers use routing algorithms to find the best route to a destination. When we say "best route," we consider parameters like the number of hops (the trip a packet takes from one router or intermediate point to another in the network), time delay and communication cost of packet transmission. Based on how routers gather information about the structure of a network and their analysis of information to specify the best route, we have two major routing algorithms: global routing algorithms and decentralized routing algorithms. In decentralized routing algorithms, each router has information about the routers it is directly connected to -- it doesn't know about every router in the network. These algorithms are also known as DV (distance vector) algorithms. In global routing algorithms, every router has complete information about all other routers in the network and the traffic status of the network. These algorithms are also known as **LS** (link state) algorithms.

- **Routing Components**

Routing involves two basic activities: determining the optimal routing paths for destination networks and

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

transporting information groups, also known as packets, through an internetwork. Within the context of routing, the latter can be referred to as switching.

- **Path Determination**

A metric is a standard of measurement, such as path length, that is used by routing algorithms to determine the optimal path to a destination. To aid in this process of path determination, routing algorithms initialize and maintain routing tables, which contain route information. This information can vary widely depending on which routing algorithm generated the routes. Routing algorithms fill routing tables with a list of networks and its corresponding "next hop" on the way its destination. When a router receives an incoming packet, it checks the destination address and attempts to associate this address with a next hop.

Algorithm Types

- Static versus dynamic
- Single-path versus multi-path
- Link state versus distance vector

- **Dynamic vs. Static**

Static routing algorithms are hardly algorithms at all, but are table mappings established by the network administrator prior to the beginning of routing. These mappings do not change unless the network administrator alters them. Algorithms that use static routes are simple to design and work well in environments where network traffic is relatively predictable and where network design is relatively simple.

Because static routing systems cannot react to network changes, they generally are considered unsuitable for today's large, changing networks. Most of the dominant routing algorithms in the 1990s are dynamic routing algorithms, which adjust to changing network circumstances by analyzing incoming routing update messages. If the message indicates that a network change has occurred, the routing software recalculates routes and sends out new routing update messages. These messages permeate the network, stimulating routers to rerun their algorithms and change their routing tables accordingly.

Dynamic routing algorithms can be supplemented with static routes where appropriate. A router of last resort (a router to which all unroutable packets are sent), for example, can be designated to act as a repository for all unroutable packets, ensuring that all messages are at least handled in some way.

- **Single-Path vs. Multipath**

Some sophisticated routing protocols support multiple paths to the same destination. Unlike single-path algorithms, these multipath algorithms permit traffic multiplexing over multiple lines. The advantages of multipath algorithms are obvious: They can provide substantially better throughput and reliability.

- **Link State vs. Distance Vector**

Link-state algorithms (also known as shortest path first algorithms) flood routing information to all nodes in the internetwork. Each router, however, sends only the portion of the routing table that describes the state of its own links. Distance- vector algorithms (also known as Bellman-Ford algorithms) call for each router to send all or some portion of its routing table, but only to its neighbors. In essence, link- state algorithms send small updates everywhere, while distance- vector algorithms send larger updates only to neighboring routers.

Because they converge more quickly, link- state algorithms are somewhat less prone to routing loops than distance- vector algorithms. On the other hand, link- state algorithms require more CPU power and memory than distance- vector algorithms. Link-state algorithms, therefore, can be more expensive to implement and support. Despite their differences, both algorithm types perform well in most circumstances.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Trust Management Architecture

Using a trust-management system for controlling security-critical services frees the application developer from a number of often difficult (and subtle) design and implementation issues and allows users to take advantage of a flexible, standard, application-independent language for specifying policy. Before trust management, every application had to provide its own mechanisms for specifying policy, interpreting credentials, and binding user authentication with the authorization to perform "dangerous" operations. Trust-management systems, on the other hand, provide a simple interface that takes care of all of these things. All the application designer has to do is identify the trust management questions in the application and formulate appropriate queries to the trust-management system.

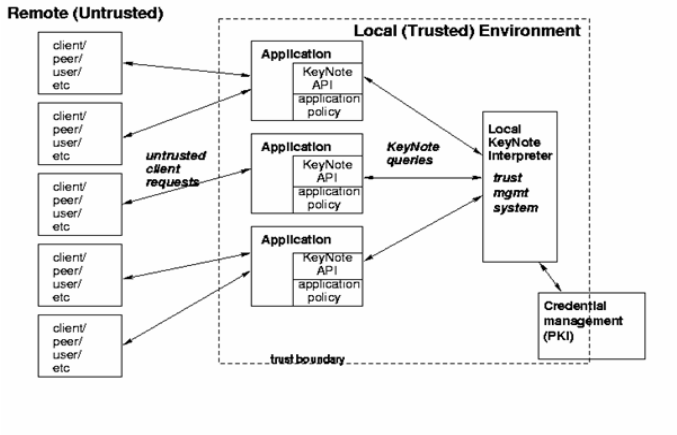


Fig. 1. Keynote trust Management Architecture

II. IMPACT OF ANT ALGORITHM ON GRID COMPUTING

Grid computing is a term used to describe both a platform and type of application. A Grid computing platform dynamically provisions, configures, reconfigures, and de provisions servers as needed. Servers in the Grid can be physical machines or virtual machines. The grid computing environment typically include other computing resources such as storage area networks (SANs), network equipment, firewall and other security devices. Grid computing also describes applications that are extended to be accessible through the Internet. These Grid applications use large data centers and powerful servers that host Web applications and Web services. Anyone with a suitable Internet connection and a standard browser can access a cloud application. The primary components of Grid architecture are:

- **Users/Brokers:** Users or brokers acting on their behalf submit service requests from anywhere in the world to the Data Center and Grid Servers to be processed.
- **SLA (Service Level Agreements) Resource Allocator:** The SLA Resource Allocator acts as the interface between the Data Center/Grid service provider and external users/brokers. It requires the interaction of the defined scheduled mechanisms to support SLA-oriented resource management.

To access the grid resources and execution, it can be divided in to three phases like resource recovery, scheduling, and executing. In the second phase find the best match between the set of jobs and available resources. The second phase is a NP-hard Problem [6]. The computational grid is a dynamic and unpredictable behavior. They are:

- Computational performance of each resource varies from time to time.
- The connection between computers and mobile phones may be unreliable.
- The resources may join or give up the grid at any time
- The resource may be occupied without a notification.

The scheduling of grid architecture is dynamic in nature and moreover Grid middleware and applications are using local scheduling and data co-scheduling. The approach of replication has been also applied and assisted in scheduling and optimization of replication. There are different existing algorithms like the Genetic algorithm (GA) is used for searching large solution space. On other hand,

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

simulated Annealing (SA) is an iterative technique that considers only one possible solution for each meta-task at a time.

ACO algorithm can be interpreted as parallel replicated Monte Carlo (MC) systems. MC systems are general stochastic simulation systems, that is, techniques performing repeated sampling experiments on the model of the system under consideration by making use of a stochastic component in the state sampling and/or transition rules. Experimental results are used to update some statistical knowledge about the problem. In turn, this knowledge can also be iteratively used to reduce the difference in the estimation of the described variables and directing the simulation process toward the most interesting state space regions. Analogously, in ACO algorithms the ants sample the problem's solution space by repeatedly applying a stochastic decision policy until a feasible solution of the considered problem is found. The sampling is realized concurrently by a collection of different instantiated replicas of the same ant type. Each ant "experiment" allows to adaptively modifying the local statistical knowledge on the problem structure. The algorithm is recursive in nature.

III. PROPOSED ALGORITHM

The classic ant colony algorithm can be described as follows:

Step 1. Initialize

Step 2. Loop /* An iteration */

Step 3. Each ant is positioned on a starting node.
Loop /* A step */

Step 4. Each ant applies a state transition rule to incrementally build a solution and a local pheromone updating rule until all ants have built a complete solution

Step 5. Global pheromone updating rule is applied until end condition.

Step 6. Stop further iterations

Each edge between node (r, s) has a distance or cost associate **8** (r, s) and a pheromone concentration **T** (r, s). The equation 1 is the state transition rule, which is a probabilistic function for each node u, which has not been visited by each placed ant on node r.

$$P_k(r,s) = \frac{[\tau(r,s)][\eta(r,s)]^{\beta}}{\Sigma[\tau(r,u)][\eta(r,u)]^{\beta}} \quad (1)$$

The parameter **P** determine the relevance of the pheromone concentration compared with the distance or cost, **T** (r, s) Global pheromone updating rule can be applied as:

$$\tau(r,s) = (1-\alpha)\tau(r,s) + \Sigma\Delta\tau_k(r,s) \quad (2)$$

Where **a** is the pheromone evaporation factor between 0 and 1 and **AT,** (r, s) is the reverse of the distance or cost done by ant k, if (r, s) is its path and is 0 if it is not in the path. The steps can be modified to manage grid architecture. The grid is visualized is the collection of clustered services, hence the live services of grid behaves like an ant, when it find its file object, the ant died. Subsequently, considering the prime component of grid computing, the *compute grid* and *storage grid* can be modeled as virtual services of grid. Every time a request is processed on a grid cluster site, **T** is updated for all the site connections and thus the "(2)" can be modified by associating a parameter **t** .

t(r,s) = (l-oc **)T(r,s)** + 2AT, **(r,s)** (3)

The dot operator represents time for each grid scheduling service. Therefore, the **a** is introduced which expresses the evaporation factor under time slot of grid service. The heuristic can be divided into two categories for grid-based services e.g. on-line mode service and the batch mode service. In online mode, whenever a request arrive, it immediately

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

allocate to the first free resource allocator. The arrival order of the request in grid is important in this method. Here, each service request is considered only once for matching and scheduling. In batch mode, the requests are collected; the scheduler considers the approximate execution time for each task and use heuristic approach to possibly make better decision. The function free [j] – return time, when the resource allocators M_j will be free. We consider,

$$free\ [j] = I_{\Delta} + ET_{ij},$$

where, I_{Δ} is the initial time slot of request of service made on the grid architecture and ET_{ij} is the execution time matrix of request r_i on resource allocator m.

The scheduling of resource allocator on the grid service proposes the probability of servicing the request:

$$P_{ij} = \frac{ph_{ij}\ \eta_{ij}\ (1/ETij)}{\sum ph_{ii}\ \eta_{ii}\ (1/ETij)}$$

Where, - η_{ij} is the attractiveness of the move as computed by heuristic information indicating a prior desirability of that move. $\mathbf{ph_{ij}}$ fast and accuracy of the grid service in the past (with lower $\boldsymbol{\alpha}$**)** to make that particular move (it represents therefore a posterior service accomplishment indication of the desirability of that request) ET_{ij} Execution Matrix of service and resource allocator. In this proposed model, we select the highest probability's 'i' and 'j' are the next request of service r_i executed on the resource allocator j.

III.CONCLUSION

This paper is the first to develop trust management architecture for Grid security solutions based on Subjective Logic. We have identified the requirements of trust management for Grid computing from security point of view. We then develop trust management architecture to meet the requirements defined. This trust management architecture is designed to be transparent to the Grid platforms. It thus can easily be instantiated in a practical application as a separate layer, and thus allows seamless integration to different Grid computing platforms. Once instantiated this architecture allows explicit trust policies to be defined and managed. In this paper, a heuristic algorithm based on modified ant colony optimization has been proposed to initiate the service load distribution under grid computing architecture. The simulation doesn't consider the fault tolerance issues. Due to absence of any restore time in service and resource allocator distribution, it is expected that continuous ant colony with other modified parameters could demonstrate better results compared to other optimization models, even in faulty service request and disrupted resource allocator.

REFERENCES

[1] F. A. Maheswaran and M., "Evolving and managing trust in grid computing systems," in *Proceedings of the 2002 IEEE Canadian Conference on Electrical Computer Engineering*, 2002.

[2] L. Shen and J., "A mission-aware behavior trust model for grid computing systems," in *Intl Workshop on Grid and Cooperative Computing (GCC2002)*, Sanya, China, Dec. 26, 2002.

[3] S. Song and K. Hwang, "Fuzzy trust integration for security enforcement in grid computing," in *International Symposium on Network and Parallel Computing(NPC2004)*, submitted March 22, 2004.

[4] C. Lin and V. Varadharajan, "Trust relationship in mobile agents - a reflexive approach," in *Proceedings of Internation Conference of Agent-Based Technologies and Systems 2003 (ATS03)*, Calgary, Canada, August 2003, pp. 81–88.

[5] A. Abdul-Rahman and S. Hailes, "Using recommendations for managing trust in distributed systems," *In Proceedings of IEEE Malaysia International Conference on Communication'97 (MICC'97), Kuala Lumpur, Malaysia*, 1997.

[6] M. R. Thompson, D. Olson, R. Cowles, S. Mullen, and M. Helm, "Ca-based trust model for grid authentication

and identity delegation," in *http://www.gridcp.es.net/Documents/GGF6/TrustModel-final.pdf*. Grid Certificate Policy WG, October 2002.

[7] J. Broch, D. Maltz, D. Johnson, Y. Hu and J. Jetcheva, "A Performance Comparison of Multi Hop Wireless AdhocNetwork Routing Protocols", Carnegie Mellon MONARCH Project, October 1998, http://www.monarch.cs.cmnu.edu/.

[8] C.E.Perkins and E.M.Royer, "Ad-Hoc On Demand Distance Vector Routing", Proceedings of the IEEE Workshop on Mobile Computing Systems and Applications WMCSA), February 1999.

[9] C.E.Perkins and E.M.Royer, "Ad-Hoc On Demand Distance Vector Routing", Proceedings of the IEEE Workshop on Mobile Computing Systems and Applications (WMCSA), February 1999.

[10] E. Bonabeau, M. Dorigo and G. Theraulaz, Swarm Intelligence: From Natural to Artificial Systems, OxfordUniversity Press, 1999.

[11] M. Dorigo and G. DiCaro, "Ant Colony Optimization: a New Meta-Heuristic", Proc. 1999 Congress on Evolutionary Computation, July 6-9, 1999, pp. 1470-1477.

Prof. R.Rameshkumar is pursuing his PhD at JNT University, Hyderabad under the guidance of Prof.Dr.A.Damodaram, Director of UGC Academic Staff College of JNT University Hyderabad. He has obtained his Bachelor Degree in Computer Science and Engineering from Mookamibigai College of Engineering (Bharathidasan University) and Master Degree in Computer Science and Engineering from Arulmigu Kalasalingam College of Engineering(M.K.University).

He joined as Faculty of Computer Science and Engineering in 1989 at JNTU, Hyderabad. He worked in the JNTU in various capacities since 1989. Presently he is a professor in Computer Science and Engineering Department. In his 19 years of service Dr. A. Damodaram assumed office as Head of the Department, Vice-Principal and presently is the Director of UGC Academic Staff College of JNT University Hyderabad. He was board of studies chairman for JNTU Computer Science and Engineering Branch (JNTUCEH) for a period of 2 years. He is a life member in various Professional bodies. He is a member in various academic councils in various Universities. He is also a UGC Nominated member in various expert/advisory committees of Universities in India. He was a member of NBA (AICTE) sectoral committee and also a member in various committees in State and Central Government

A Forager Bees Behaviour Inspired approach to predict the forthcoming navigation pattern of online users

V.Mohanraj
Assistant Professor/IT
Sona College of Technology
Salem, Tamilnadu, INDIA
bvmohanraj@yahoo.in

Dr.R.Lakshmipathi
Professor/EEE
St.Peters Engineering College (Deemed University)
Chennai, Tamilnadu, INDIA
drrlakshmipathi@yahoo.com

J Senthilkumar
Assistant Professor/IT
Sona College of Technology
Salem, Tamilnadu, INDIA
rajkrishcounty@gmail.com

Y.Suresh
Assistant Professor/IT
Sona College of Technology
Salem, Tamilnadu, INDIA
shuresh_22@yahoo.co.in

***Abstract*— The World Wide Web is continuously growing and become de facto place to conduct online business. In the current internet world, peoples are more attracted towards participating in the e-commerce sites. The real challenge for the web master of any such website is to find the users need in advance and provide the resources pages that keep them interested in browsing their site. It is easy for any unsatisfied user to reach out the counterpart site in a single click. Many Web usage mining methods were adopted to work on web server log and predict the forthcoming navigation pattern of user. However, the accuracy of the methods can't satisfy the user especially in huge site.**

This paper presents the forager bees behaviour inspired Forager agent based architecture that uses its collective intelligence for predicting the forthcoming navigation pattern of user. Our practical implementation shows that accuracy and coverage measures are very much improved than existing methods.

Keywords-Web Usage Mining; Web Personalization; Artificial Bee Colony.

I. INTRODUCTION

In the current internet world, there are many e-commerce sites compete with each other to attract the user. It becomes mandatory for web master to predict the future navigation of user and recommend those to users. This makes the user to browse the site with lot of satisfaction. In case of unhappiness, it's easy for online user to switch over to another e-commerce site that provides the same kind of service.

All the e-commerce sites are focusing on how to provide the excellent personalized access to users on their sites. The solution of the problem is web usage mining (WUM). WUM [8] is part of web mining which deals with the extraction of knowledge from server log file [4] [5].Source data mainly consists of the logs, that are collected when user access web server and might be represented in standard format. WUM has become very critical for website management, creating adaptive website, business and support services, personalization and network traffic flow analysis.

Typically, the WUM based forthcoming navigation pattern capturing process can be divided into FrontEnd and BackEnd with respect to the web server activity [6].The activities of BackEnd component is focused on building the knowledge base by analyzing server log file which records user web usage data. The activities of the FrontEnd component are classifying the current user navigations to any one of cluster formed in BackEnd Phase and infer the useful pattern to predict the future navigations of user. A particular feature of our paper is that achieving the accuracy of excellence in predicting the forthcoming navigation pattern of user using the collective intelligence of Artificial Bee System [15]. This system is relatively new member of swarm intelligence. It tries to model natural behaviour of real honey bee in food foraging. Honey bee use several mechanisms like waggle dance, round dance and tremble dance to exchange the information about location and profitability of food source. This makes them a good candidate for developing new intelligent search algorithms. Artificial Bee system has three area of study: Foraging Behaviour, Marriage Behaviour and Queen Bee concept. Our paper focuses on the usage of foraging behaviour of bee in the FrontEnd Phase of Forager architecture to achieve excellence in capturing the forthcoming browsing pattern of online users.

The paper is organized as follows. In section II, we review the related automatic recommendation systems and reported as literature survey. In section III describes the different phases of Forager agent based architecture and focuses on the Greatest Common Subsequence detection algorithm which is used by foragers and Intuition Deductive inference engine used by the onlooker in the FrontEnd phase of the architecture. In section IV, the illustration of Forager Agent system is explained. In section V, the results of our practical implementation are reported. Finally, section VI concludes our work.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

II. RELATED WORKS

Recently, a number of studies have been proposed to capture the forthcoming navigation pattern of web users. We have conducted investigation on different WUM system [10] and architecture that can be matched with our proposed system.

Analog [1] is one of the first web usage mining systems. It consist of online and offline component. The offline component builds session clusters by analyzing user navigation pattern recorded in the log file. In the online component part, active user session is classified according to the generated model. The classification allows identifying the pages matches with the active session and returning the requested page with list of suggestion. Clustering approach of system affected by several limitation especially scalability and accuracy. There is variety of clustering algorithms available for usage. Each approach could have different type of cluster [Exclusive (K-Means), Overlapping (Fuzzy C Means) and Hierarchical].It's difficult to compare the performance of algorithm on large dataset like web log. In addition, Clustering approach used in all recommender system needs to be back up by excellent classification method. Analog did not have the proper classification approach over the overlapping cluster.

A Web personalizer system [11] provides dynamic recommendation, as a list of hyperlinks to users. In the Web personalizer system, analysis is based on the usage data combined with structure formed by the hyperlinks of site. Aggregated usage profile is obtained by applying data mining technology [i.e clustering, association rule] in pre-processing phase [12]. In this phase web server logs are converted into cluster made up of set of pages with the common usage characteristic. The online phase considers the active user session in order to find match among user activities and discover usage profile. Matching entries are then used to compute a set of recommendations which will be inserted into the last requested page as a list of hypertext links. Webpersonalizer is good example two tier architecture for personalized system. However the accuracy of the Webpersonalizer is affected by association rule mining [16] used for discovery of frequent item set in web log data. The main problem with the association rule mining method is discovery of contradictory association rules. As a result of inconsistent rules, predicting accuracy of system is degraded. Even the non redundant association rule mining algorithm does not help the system because of the web log data nature where the number of page hits is high.

Another WUM system called SUGGEST [2] provides useful information to optimize web server performance and make easier the web user navigation .SUGGEST adopts a two level architecture composed by an offline creation of historical knowledge and online component that understands the users behaviour .SUGGEST uses the markov model for calculating the probability of a page the web user visit in future after visiting pages in the same session. This system uses the high order markov model [7] to improve the accuracy. However, the system can't be used for web site made up of large number of pages due to high space complexity. The limitation of system might be a) the memory required to store web server pages is quadratic in the number of pages. It is a sever limitation in the huge web site. b) SUGGEST does not permit us to manage web site made up of pages dynamically generated.

Our survey reveals that there is a race for finding architecture [3][13] and classification algorithm to improve the accuracy of capturing forthcoming navigation pattern of online users. But still the accuracy does not meet the satisfaction. In our work, we propose advanced Forager agent based architecture and novel user navigation classification approach in the architecture for improving accuracy and space complexity. Our Forager agent based architecture is the inspiration from artificial bee colony introduced by [14][15].

1. Each employed bee determines a food source, which is also representative of a site, within the neighbourhood of the food source in its memory and evaluates its profitability.
2. Each employed bee shares its food source information with onlookers waiting in the hive and then each onlooker selects a food source site depending on the information taken from employed bees. Each onlooker determines a food source within the selected site by herself and evaluates its profitability using the collective intelligence.
3. Employed bees whose sources have been abandoned become scout and start to search a new food source randomly (Fluctuation).

The step (2) of the algorithm is implemented in the FrontEnd phase of Forager agent architecture. In the FrontEnd Phase, fleet of forager is originated by onlooker agent on number of clusters formed in the BackEnd Phase. Each forager search the web pages in the cluster based on reinforcement given from onlooker agent. Each forager executes the Greatest Common Subsequence detection algorithm on its cluster of web pages and also runs the scoring algorithm. The final score of each forager is received by onlooker agent. In case of more than one profitable cluster, it is the onlooker agent that runs the intuition deductive inference engine to choose the best one among the alternatives. Finally, recommends the selected one as forthcoming browsing pattern of the user.

III. DIFFERENT COMPONENTS OF TWO TIER FORAGER AGENT BASED ARCHITECHUTRE

According to different functionality, our proposed architecture can be divided into two main phase Back End and Front End. Both these phases are tightly coupled with each other and work closely together. The Figure 1 and 2 depicts the BackEnd and FrontEnd architecture of two tier Intelligent Forager Agent respectively. In the Back End Phase there are two main module, Data pre-processing and user navigation mining. The main modules of the Front End phase are onlooker agent, forager agent and Intuition deductive inference engine.

A. BackEnd Phase of Architecture

Two main major modules of the BackEnd Phase are Data pre-processing and user navigation pattern mining. In

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

this phase, we perform Data pre-processing on server log to capture navigation session and after that we apply algorithm to mine user navigation pattern. The Detailed module of BackEnd Phase is shown in Figure 3.

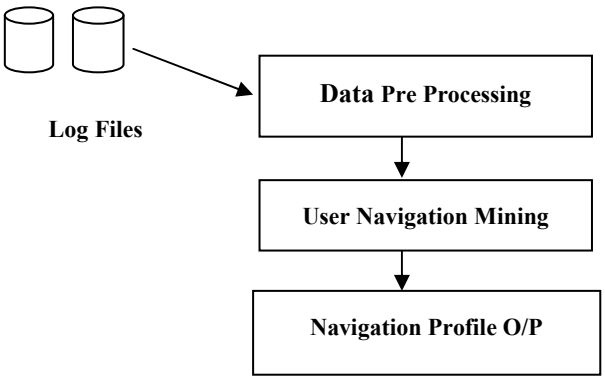


Figure 1. BackEnd Phase of Two Tier Forager Agent Architecture

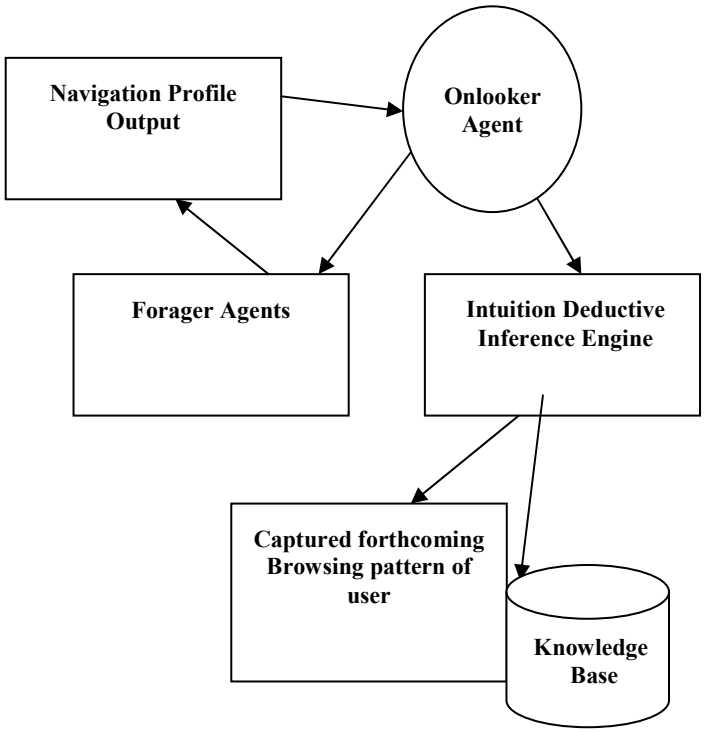


Figure 2. FrontEnd Phase of Two Tier Forager Agent Architecture

a) Data Pre-processing

The pre-processing of web logs is usually complex and time demanding. It comprises of four different tasks **1) Data collection:** A flat file was constructed from original weblog file. Each record of the file consists of time, ip address, name, requested resource (URL) and HTTP Status code. **2) Data Cleaning:** In this step, we perform the removal of all the data tracked in web log that are useless for mining purpose such as Navigation session performed by robots and web spider. **3) Session Identification and reconstruction:** it involves i) Identifying the different users session from usually very poor information available in log files and ii) Reconstructing the user's navigation path within the identified session. 4**) Content and Structure Retrieving:** Mostly all WUM uses the visited URL as the poor source of information. They do not convey any information about actual page. We employ the content based information to enrich web log data in the form of maintaining ontology for each web page in the Array data structure where each index corresponds to page number of web site. This information will be used by the Intuition deductive engine of Front End for the choosing the best classification.

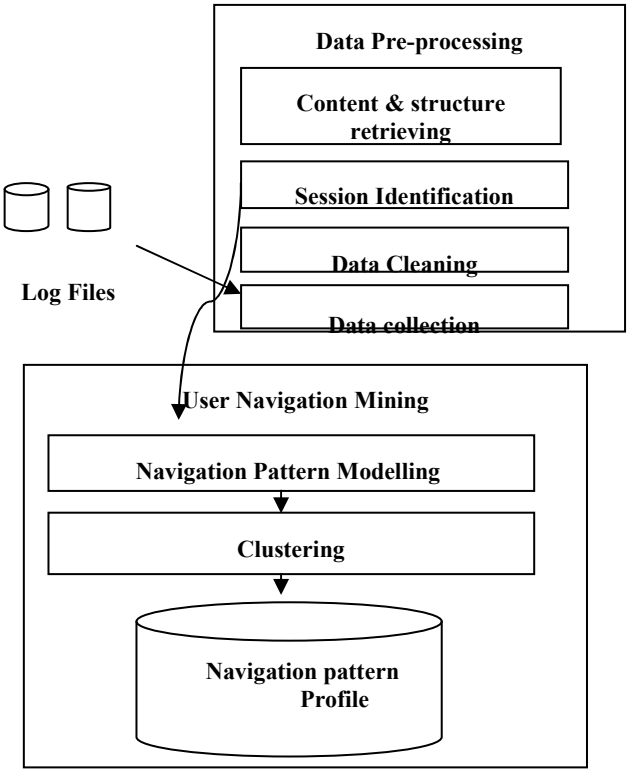


Figure 3. Components of BackEnd

b) User Navigation Mining

After the data pre-processing, we perform a navigation pattern mining on the identified session. We perform clustering which aims to group session into clusters based on their sharable properties. These patterns will be further used to facilitate the user profiling process of the system. It includes two main modules:

1) Navigation Pattern Modeling

In this step, web pages accessed are modeled as undirected Graph G= (V, E). The set Vertex (V) identifies the different web pages hosted on the web server model. The edge weights are determined by the following equation

$$WPij = \frac{Cij}{Max\ \{Ci\ , Cj\ \}} \qquad (1)$$

Where C_{ij} is the number of session containing both pages i and j. C_i and C_j are respectively the number of sessions containing only pages i or page j. Dividing by the maximum between single occurrences of two pages has the effect of reducing the relative importance of links involving index pages. Such pages are those that generally do not contain

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

useful content and are used only a starting point for a browsing session. The edge weights (WP_{ij}) are kept in the adjacency matrix WA where each entry WA_{ij} contains the value computed according to equation (1). To limit the number of edge in such graph, elements of WA_{ij} whose value is less than a threshold are known to be less correlated and thus discarded.

2) Clustering

We apply a graph partitioning algorithm to find groups of strongly correlated pages by partitioning the graph according to its connected components. Clusters are formed by starting from a Vertex a DFS on the graph induced by WM is applied to search for the connected component reachable from this vertex. Once the component has been found, the algorithm checks if there are any nodes not considered in the visit. If it so, it means that a previously connected has been split and therefore, it needs to be identified. To do this, DFS is again performed by starting from one of the nodes not visited. In the worst case, when the entire URL in the same cluster, the cost of the algorithm will be linear in the number of edges of the complete graph G. Before the clusters are put into navigational pattern profile, the clusters are ranked based on values store in the WM matrix. It will be used for classification performed by foragers based on the Greatest Common Subsequence Detection algorithm and also used for knowledge eliciting.

B. FrontEnd Phase of Architecture

In the FrontEnd Phase of our system, URL request of the user is processed by the Onlooker Agent (OA) and Forager Agents (FA). In the case of multiple profitable outputs, the best option is choosen by the Intuition Deductive Inference Engine (IDIE) which is run by Onlooker Agent. Finally, captured imminent browsing pattern is suggested to the user who initiated the URL request to web server.

a) Work of Onlooker and Forager agent in FrontEnd Phase

The critical component of our system is Onlooker agent and Forager agent. Main inputs for these agents are

1. Navigation pattern profile: It consists of clusters formed in the BackEnd Phase of our system.

2. Live session window: A Sequence $LSW=\{lwp_1,lwp_2,....,lwp_m\}$ is the current size of live session window where m is the size of the current active session window.

On receiving URL's in the form of Live Session Window, Onlooker initiates the Foragers that correspond to number of clusters in the navigation profile. Each **Forager agent (FA)** acts on the Navigation profiles by executing the novel Greatest Common Subsequence Detection to discover the subsequence which may be considered as imminent browsing pattern of web user. Each FA submits the profitable score along the discovered subsequence to the onlooker.

When the **Onlooker Agent (OA)** receives the profitable score along the sub sequences, it starts to decide the best profitable source of navigation profile. In the case of close race between the sub sequences, **Intuition deductive inference engine (IDIE)** plays a crucial role in selecting the right cluster. The main objective of IDIE is to test the each cluster of discovered navigation pattern against already built knowledge base and choose the cluster which attains the maximum number of matches with the knowledge base. Finally, Onlooker suggests the output of IDIE as the imminent browsing pattern of web user. The cooperation between Onlooker Agent (OA) and Forager Agents (FA) is depicted in the Figure 4. The OA listens to FA. It is just similar to honey bee dancing area where Onlooker listens to dances of different foragers about the profitable food source.

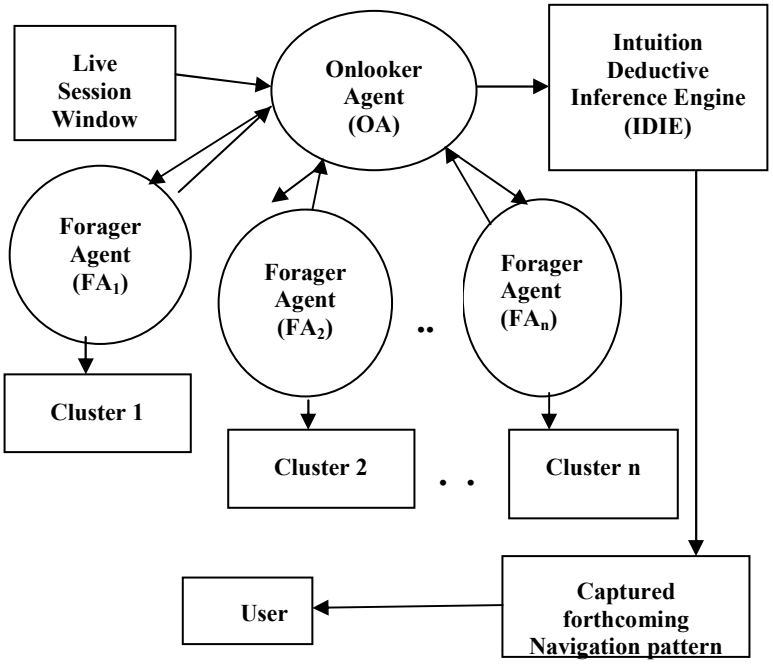


Figure 4. Onlooker Agent and Forager Agents

b) Algorithm for predicting forthcoming navigation pattern of online users

The following **Algorithm 1** depicts the working behaviour of onlooker and forager agents to capture the forthcoming browsing pattern of user in the FrontEnd phase of the Architecture.

Algorithm 1

1. Live Session Window (LSW) is given as input to the Onlooker Agent.LSW is the set of web page visited by user in the live session. LSW is represented as {LWP_1, $LWP_{2,..}$, LWP_n) where 'n' is the size of session window.
2. Onlooker Agent (OA) initiates the number of Forager Agent (FA_i) that corresponds to each cluster in Navigation Profile (NP_n) where'n' is number of cluster.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

FA_i - Forager Agent works on i^{th} cluster in NP and i Є n.

3. Initially onlooker agent assigns the arbitrary profitable score of value 100 to each **Forager Agent (FA_i)**. The score is denoted as **score(FA_i)**.This score is updated by the respective Forager Agent on the discovery of subsequence in the step 4.
4. For each Forager Agent on its assigned cluster of Navigation Profile does the following
   i. Each FAi executes the Greatest Common subsequence Detection (described in c) on its cluster in respect to **live session window (LSW)** which produces the highest degree of GCD as the discovered subsequence which could be the candidate of online user's imminent browsing pattern.
   ii. Discovered sequence is denoted as IBP = {IBP_1, IBP2,.., IBPn}.Each Forager Agent updates its initial score using Equation (2) and with the help of adjacency matrix WA built in the BackEnd Phase where each entry WA_{ij} contains the value computed according to equation (1).

$$\Delta score(\mathrm{FAi}) = \mathrm{score(FAi)} + \sum_{i=1}^{n}\sum_{j=1}^{m} WA_{\mathrm{IBPi\ LWPj}} \quad (2)$$

Where $WA_{\mathrm{IBPi\ LWPj}}$ = Value in Adjacency matrix between the each page in Imminent Browsing Pattern (IBP) discovered by Forager and pages in the Live Session Window (LSW).

   iii. After each FAi executed the steps i and ii, Forager Agent sends its updated score and discovered Subsequence to the Onlooker Agent (OA).
5. After receiving the profitable scores from each Forager agents ($FA_{i)}$, It selects the first 3 High scored Forager Agent's output. The scores are denoted as PS_1, PS2 and PS_3.

   i. Onlooker Agent computes the absolute difference between the PS_1, PS2 and PS_3 to find the closeness.

IF $|PS_1 - PS_2|$ **or** $|PS_1 - PS_3|$ **or** $|PS_1 - PS_2| \leq \beta$ **Where** β **is Uncertain Profitable Threshold value then**

There is race between discovered subsequence of PS_1 or PS2 or PS_3 to become an imminent browsing pattern of web user. Onlooker Agent (OA) sends the cluster of competing sub sequences to the Intuition Deductive Inference Engine (IDIE).The main objective of IDIE (described later) is to test the each clusters navigation pattern against already built knowledge base and choose the cluster which attains the maximum number of matches with the knowledge base.

ELSE

Onlooker chooses the PS1's sequence as best discovered subsequence.

6. Finally, **Onlooker Agent** reports the predicted forthcoming browsing pattern as PS_1 to user Or best Sequence selected by the IDIE in the case of competition.
7. Suppose, if the next user activity in live session window different from the suggested captured list then the system has to restart once again to classify the new user activities.

c) Greatest Common Subsequence Detection

Every Forager Agent initiated by Onlooker Agent should perform the similarity comparison between set of pages in Live Session Window and web pages in the cluster to discover the subsequence that could be the forthcoming browsing pattern of online user. It's clear that every forager agent has to perform some kind of pattern matching.

In the pattern matching [9], comparing the similarity between the two sequences $\vec{a}$ and $\vec{b}$ are fundamental problem. One of the fundamental problem is to determine the Greatest Common Subsequence (GCS) between $\vec{a}$ and $\vec{b}$. The GCS is a String comparison metric that measures the subsequence of maximum length common to both the sequences. Main objective of Forager Agent (FA) is to find the Greatest Common Subsequence among the sequence of paths in the form of page visits A= {a1,a2,a3,a4,.....an) , B={b1,b2,b3,b4,...bn).

Theorem 1 Let a= {a1,a2,a3,a4,.....an) and b={b1,b2,b3,b4,...bm) be the sequences and Let d={d1,d2,d3,...dn} be any GCD of a and b.

1. If $a_n=b_m$ then $d_l=a_n=b_m$ and d_{n-1} is a GCD of a_{n-1} and b_{n-1}.
2. If $a_n \neq b_m$ then $d_l \neq a_n$ implies d is a GCD of a_{n-1} and b.
3. If $a_n \neq b_m$ then $d_l \neq b_m$ implies dn is a GCD of a and b_{m-1}.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

We have implemented the GCD with added module that outputs the subsequence of indices of the two sequences that match in getting the Greatest Common subsequence. For example, If A={wp1,wp2,wp3,wp2,wp4,wp2,wp1} and B={wp2,wp4,wp3,wp1,wp2,wp1}. Their GCD is GCD= {wp2,wp4,wp2,wp1).

d) Combined Effort of Greatest Common subsequence Detection and Intuition Deductive Inference Engine (IDIE)

After performing the clustering algorithm discussed in A.b.2 of BackEnd Phase, We have a set of cluster wnp = $\{wnp_1,wnp_2,wnp_3,...,wnp_n\}$ where $wnp_i=\{wp_1,wp_2,...,wp_k\}$ is a set of K pages as a users navigational pattern for each $1\leq i \leq n$. Here, wnp is set of web navigational pattern in the cluster.

We used the navigation profile which has web navigational pattern of different cluster as facts for building the knowledge base of Intuition deductive Inference Engine. In addition, OA an ontology array data structure built in the Content retrieving step of BackEnd phase is used to form the meaningful facts. We used the unique ontological term for each web page in the site. We get the ontology term of each page from Meta tag of each page which conveys the actual content of the page. Also, Structure of Website is used to input the additional facts for knowledge base. In our WUM system, Knowledge base building is considered as critical part.

A Sequence LSW=$\{wp_1,wp_2,....,wp_m\}$ is the current size of live session window where m is the size of the current active session window. Before **Onlooker Agent** initiates the Forager Agents to execute GCD on its cluster, we need to order the lsw sequence based on their adjacency weight matrix (WM) constructed in the navigation pattern modeling. Also, we rank all the clusters based on their weight values. Each cluster weight is computed as sum of all its edges weight. After this step, each Forager Agent initiated by OA with the arbitrary profitable score applies the Greatest Common Subsequence Detection on the assigned cluster in respect to **live session window (LSW)** which produces the highest degree of GCD. Each Forager Agent sends its updated score and discovered subsequence to the Onlooker Agent (OA).

After receiving the profitable scores from each Forager agents ($FA_{i)}$, OA selects the first 3 High scored Forager Agent's output and finds whether absolute difference between them lesser then β **(Uncertain Profitable Threshold value).** In the case of competition between discovered sequences by forager agents, **Intuition deductive inference engine (IDIE)** plays a crucial role in selecting the right cluster among various options.

The main objective of IDIE is to test the each clusters navigation pattern against already built knowledge base and choose the cluster which attains the maximum number of matches with the knowledge base. The rules of knowledge base are written in InterProlog notation. Our inference engine runs on the facts to check how many of web page navigational sequence in each cluster are semantically correct. The rule base is written in such a way that it checks out each cluster against the knowledge base to find the semantic valid matching count. Intuition Deductive Inference Engine is designed to infer the facts by following Top Down inference mechanism. The excerpts of our IDIE knowledge base are shown in Table I.The fact **Next (wp_7(cse research),wp_6(research))** says that wp_7 is the semantically the next page of wp6 and values in the bracket are ontological terms gathered from OM array. The rule (1) can directly apply on the facts where as rule (2) is recursive nature.

As we discussed earlier, clusters of competing discovered sequences selected by **Onlooker Agent (OA)** is given as input to IDIE. The IDIE infers the count for each cluster that shows how many sub sequences in each cluster are semantically matching with the knowledge base. Finally, IDIE reports a cluster of maximum valid match count. This is used for preparing the intuition captured list and provided to the user as a recommendation. Suppose, if the next user activity in live session window different from the suggested captured list then onlooker agent has to restart the algorithm once again to identify the forthcoming user activities on their site.

TABLE I. EXCERPTS OF IDIE KNOWLEDGE BASE

Facts
Next (wp_7(cse research),wp_6(research))
Next (wp_{10}(it research),wp_6(research))
Next (wp_{17}(cse),wp_{24}(course))
Next (wp_7(cse research),wp_{17}(cse))
Next (wp_{27}(cse staff details),wp_{17}(cse))
.....
Rules:
Subsequence(x,y):- Next(x,y) (1)
SuperSubsequence(x,y):- Next(x,z) , Next(z,y) 2).....

IV. ILLUSTRATION OF FORAGER AGENT BASED SYSTEM

For our illustration, consider the navigation profile of BackEnd Phase as shown in Table II. Assume the Live Session Window size as 3 and set of web pages visited by user in the live session as LSW= $\{wp_{37},wp_{27},wp_{18}\}$. As stated in our algorithm 1, LSW is given as input to the Onlooker Agent (OA). OA initialize FA_i (FA_1, FA_2, FA_3, FA_4 and FA_5) that works respectively on Navigation Profile (NP_1, NP2, NP3, NP4 and NP_5) with the initial arbitrary profitable score of 100.Each Forager Agent (FA_i) executes the Greatest Common subsequence (GCD) on its own controlled cluster in respect to Live Session Window (LSW).

While Forager Agent discovers the subsequence that could become Imminent Browsing Pattern (IBP), update its score according to equation (2) i.e. Sum of adjacency matrix values for pages between LSW and IBP is added to initial profitable score. Consider the Table III that depicts the score of each Forager Agent (Score (FA_i)) and its discovered sequence (IBP_i).Each Forager Agents (FA_i) sends its Final Profitable Score (FA_i) and Discovered Sequence (IBP_i) to Onlooker Agent (OA). After receiving the profitable score from all initialized FA_i, OA selects the first three high scored FA_i.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

In this example, only FA_3 and FA_5 produced the updated profitable score. The remaining Forager Agents were not updated their initial score.

TABLE II. NAVIGATION PROFILE OF BACKEND PHASE

Navigation Profile	Clustered Navigation Pattern
NP_1	$\{wp_2, wp_{10}, wp_{15}, wp_{20}, wp_8\}$
NP_2	$\{wp3, wp_{27}, wp_{54}, wp_{100}, wp_{121}\}$
NP_3	$\{wp_2, wp_{19}, wp_{37}, wp_{27}, wp_{30}, wp_{18}, wp_{60}\}$
NP_4	$\{wp_5, wp_{15}, wp_{23}\}$
NP_5	$\{wp_7, wp_{37}, wp_{31}, wp_{27}, wp_{29}, wp_{26}, wp_{18}\}$

TABLE III. SCORE(FAi) AND ITS DISCOVERED SEQUENCE (IBPi) IN RESPECT TO LIVE SESSION WINDOW

Forager Agents (FA_i)	Initial Score	Navigation Profile (NPi)	Discovered Sequence (IBP_i)	Final Profitable Score (FA_i)
FA_1	100	NP_1	NO	100
FA_2	100	NP_2	NO	100
FA_3	**100**	**NP_3**	**$\{wp_{19}, wp_{30}, wp_{60}\}$**	**172**
FA_4	100	NP_4	NO	100
FA_5	**100**	**NP_5**	**$\{wp_7, wp_{31}, wp_2, wp_{26}\}$**	**213**

NO- No Output.

In the next step, Onlooker Agent (OA) finds the absolute difference between $|Score(FA_3) - Score(FA_5)|$ and compares the value with the β **(**Uncertain Profitable Threshold value**).** Here the value of β is assigned to be 50**.** In this example, the value of $|Score(FA_3) - Score(FA_5)|$ is 41 which lesser than β**.** This situation is the typical case of competition of who to become the Imminent Browsing pattern of user between IBP_3 and IBP_5 which are discovered sequence of FA_3 and FA_5. Now, Onlooker Agent uses the Intuition Deductive Inference Engine (IDIE) to choose the best one among the alternatives IBP_3 and IBP_5.

Onlooker Agent feds the clusters of competing discovered sequence NP_3 and NP_5 to the IDIE. The main function of IDIE is to check each page of clusters with knowledge base and find whether they are semantically valid next page in respect to pages in Live Session Window. Onlooker Agent chooses the cluster which attains maximum valid match count in respect to LSW. In this example, FA_5 is reported by the IDIE for attaining the more number of valid matches with knowledge base than FA_3. From the FA_5, OA suggest the following list as perfect Imminent Browsing Pattern of user

IBP = $\{wp_7, wp_{31}, wp_{29}, wp_{26}\}$.

V. DATA ANALYSIS REPORT

Our proposed Forager Agent based system was tested on the web log dataset of Sona College. This site was deployed in the IBM's Web sphere Application Server. All the algorithms of BackEnd and FrontEnd phase were implemented in the JAVA. The knowledge base used in the IDIE of FrontEnd phase was implemented using the InterProlog. InterProlog provides us the ability to call prolog goal through a prolog object and for prolog to invoke any JAVA method through a Java Message Predicate. The BackEnd phase of our system was tested on the weblog entries of 120 students over a period of 8 weeks. There were approximately 52,745 entries in the log file. For our input dataset, BackEnd phase had produced the navigation profile output which consists of 15 clusters. The performance of the proposed system was analyzed based on the two metrics. They are namely Accuracy and Coverage.

a) Accuracy based analysis of Forager Agent based System

Accuracy measure is defined as a degree to which captured imminent browsing pattern as suggested by the system matches with the actual browsing pattern of user. It is given by

$$\text{Accuracy} = \frac{|P(IBP_{np}, LSW) \cap Original_{np}|}{|P(IBP_{np}, LSW)|} \quad (3)$$

LSW - Live Session Window $P(IBP_{np}, LSW)$ - Navigation Pattern in predicted imminent browsing pattern of user. $Original_{np}$ - Original Navigational pattern of user

The Figure 5 depicts the accuracy of our Forager Agent based system as Live Session Window (LSW) size is increased. Our results show that increase of LSW size gives more wisdom to system that improves the accuracy.

The Figure 6 depicts the comparison of our Forager Agent based system with the other two recommendation system namely Web personalizer and SUGGEST. Our results show that Our Forager Agent based system outperforms the recommendation system with the excellent behaviour of Forager Agents which uses Greatest Common Subsequence Detection to predict the forthcoming browsing pattern of user. In the case of Competition between alternative, Onlooker Agent uses the Intuition Deductive Inference Engine to choose the best one and thereby completely avoids the misclassification in finding the forthcoming browsing pattern compared with other system. Our Forager Agent based system outperforms the other recommendation system and achieves the accuracy of 92% compare to the Web personalizer (81%) and SUGGEST (83%) when Live size window is 10.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

b) Coverage based analysis of Forager Agent based System

Coverage measure is defined as the ability of two tier Forager agent based system to produce all page views that are most likely visited by the user. It is given by

$$\text{Coverage} = \frac{\left| P(IBP_{np}, LSW) \bigcap Original_{np} \right|}{\left| Original_{np} \right|} \quad (4)$$

The Figure 7 depicts the coverage of our Forager Agent based system as Live Session Window (LSW) size is increased. The practical implementation of our proposed two tier Forager agent based system on sona college dataset prove that there is increase in the accuracy of predicting the navigation pattern of online user and also Coverage measure is excellent.

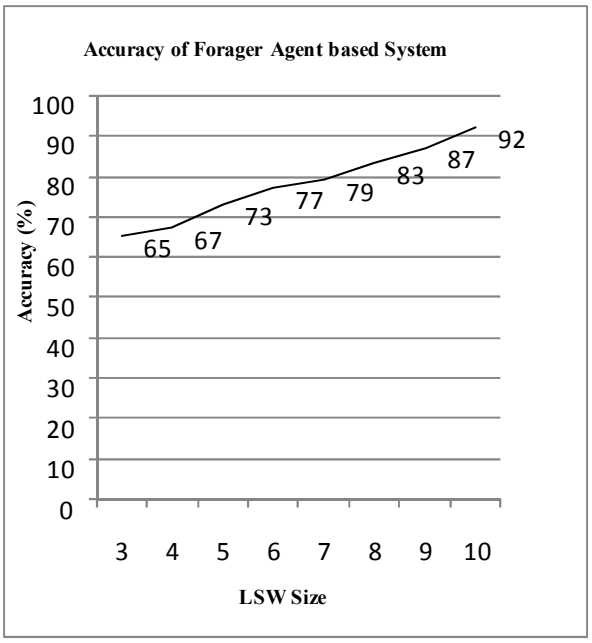


Figure 5. Accuracy of Forager Agent Systems

VI. CONCLUSIONS

Our proposed two tier Forager agent based system presented in this paper was inspired from onlooker bee making a decision of profitable food source using a Collective intelligence of Foraging Behaviour in Bee's Hive. In our system, fleet of forager is originated by onlooker agent on number of clusters formed in the BackEnd Phase. Each forager executes the Greatest Common Subsequence detection on its cluster of web pages and also runs the scoring algorithm to discover the subsequence that could be the imminent browsing pattern of user. Each forager agent sends the final score to onlooker agent. In case of more than one profitable cluster, it is the onlooker agent that runs the intuition deductive inference engine to choose the best one among the alternatives. Finally, recommends the selected one as forthcoming navigation pattern of the user. The practical implementation shows that our approach really improves the accuracy of predicting the forthcoming browsing pattern of user and satisfies the users compared to other system.

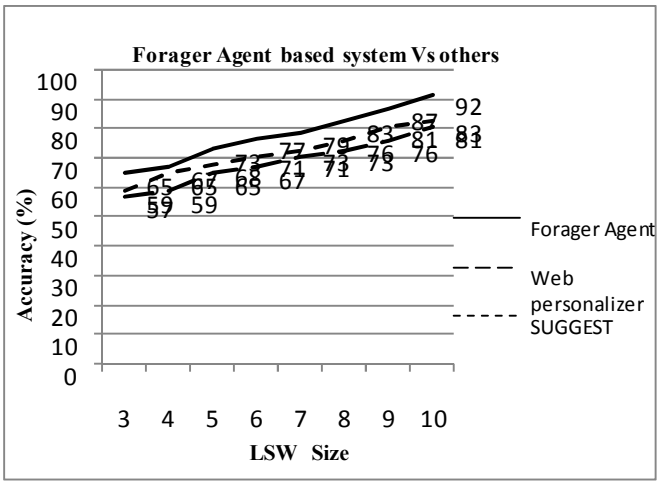


Figure 6. Accuracy of Forager Agent based System Vs Other Systems

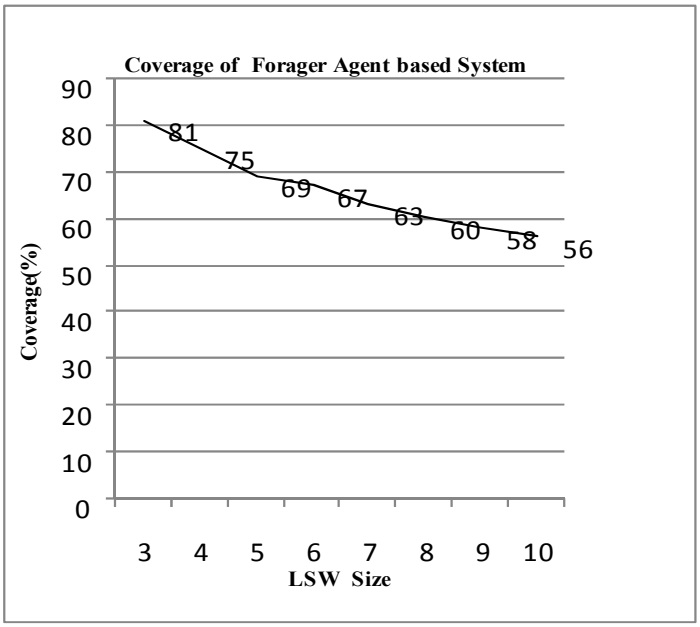


Figure 7. Coverage of Forager Agent based System

REFERENCES

[1] R. Agrawal and R. Srikant, "Mining sequential patterns", International Conference on Data Engineering (IDCE), 1996, Taiwan, pp.3-11.

[2] R. Baraglia and Palmerini, "SUGGEST: A Web usage mining system", Proceeding of International Conference on Information Technology: Coding and Computing, 2002, pp.282-287.

[3] R. Baraglia and F. Silversti,"Dynamic Personalization of Web Sites without User Intervention", Communication of the ACM, 2007, pp.63-67.

[4] R. Baragila and F. Silvertri,"An online recommendation system for large web sites", Web Intelligence,IEEE/WIC/ACM,2004,pp.20-24.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

[5] F.H. Chanchary, I. Haque and Md.Khalid,"Web Usage Mining to Evaluate the Transfer of Learning in a Web based Learning Environment", Knowledge Discovery and Data Mining of IEEE, 2008, pp.249-253.

[6] R. Cooley, J. Srivastav and B.Mobasher,"Automatic Personalization based on Web Usage Mining", Communication of the ACM,2000, Volume 43,issue 8,pp.142-151.

[7] Deshpande and G. Karypis,"Selective Markov models for predicting web page access", Transactions on Internet Technology, 2004, Vol.4, No.2, pp.163-184.

[8] E. Frias-Martinez, V. Karamcheti,"Reduction of user perceived latency for a dynamic and personalized site using web mining technique", WebKDD,2004,pp.12-22.

[9] D.S. Hirschbergand and J.D. Aho.Ullman,"Bounds on the Complexity of the longest Common subsequence problem", J.Assoc. Comput. Mach. ACM,1976, pp.1-12.

[10] R. Liu and V. Keselij,"Combined mining of Web Server logs and Web contents for Classifying user navigation pattern and predicting users future requests", Data & Knowledge Engineering,Elsevier,2008,pp.304-330.

[11] B. Mobasher,"Web personalizer: A Server Side Recommender System Based on Web Usage Mining", 1991, In. Technical Report TR-01-004.

[12] Nakagawa and B. Mobasher,"A hybrid web personalization Model based on site connectivity", WebKDD,2003, pp.59-70.

[13] Jalali, N. Mustapha, A. Mamat and N. Sulaiman.Md "OPWUMP -An Architecture for online Predicting in WUM-based Personalization system", In 13th International CSI Computer Science, 2008 Springer Verlag.

[14] D. Karaboga and B. Basturk ," On the Performance of Artificial Bee Colony (ABC) Algorithm, Applied Soft Computing,2008, Volume 8, Issue 1, Pages 687-697.

[15] V. Tereshko, A. Loengarov,"Collective decision-making in honey Bee foraging dynamics, Comput. Inf. Syst. J., 2005, pp. 1352-1372.

[16] M. Yan, H. Jacobsen, Garcia-Molina, and U. Dayal, "From User Access Patterns to Dynamic Hypertext Linking," Comp. Networks and ISDN Sys, 1996, vol. 28, pp. 1007-14.

AUTHORS PROFILE

V.Mohanraj received his ME Computer science and Engineering from Anna University, Chennai in 2004. He is currently working as Assistant professor in IT department of Sona College of Technology. He is pursuing PhD degree in Anna University, Chennai. His research area includes web mining, database and intelligent system.

Dr.R Lakshmipathi received his BE from College of Engineering, Anna University, India, in 1971, the ME and PhD in Electrical Engineering from College of Engineering, Anna University, India in 1973 and Indian Institute of Technology (IIT), Chennai in 1979, respectively. He has 36 years of teaching experience at UG degree level out of which 10 years in PG degree level. He worked as principal in Govt. College of Engineering and held the various prestigious posts like Dean, Regional Research Director , Chairman for board of BE exams, Member of Academic auditing committee, AICTE, University and State Govt. expert committee member. He is currently a professor of Electrical Engineering, St. Peters Engineering College (Deemed University), Tamilnadu. His research interest includes Electrical power semi conductor drives, Signal processing and Web Mining.

J.Senthilkumar received his ME Applied Electronics from Anna University, Chennai in 2004. He is working as assistant professor in IT Department of Sona College of Technology and pursuing PhD degree in Anna University, Chennai. He is active in the research area of data mining and Mobile computing.

Y.Suresh received his ME Applied Electronics from Anna University, Chennai in 2004. He is working as assistant professor in IT Department of Sona College of Technology and pursuing PhD degree in Anna University, Chennai. He is active in the research area of data mining, control system and Mobile computing.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Quality of Service Issues in Wireless Ad Hoc Network (IEEE 802.11B)

Mohammed Ali Hussain[1], Mohammed Mastan[2], Syed Umar[3]
[1]Research Scholar, Dept.of CSE, Acharya Nagarjuna University, Guntur, A.P., India.
hussain_ma2k@yahoo.co.in
[2]Research Scholar, Dept.of CSE, JNT University, Kakinada, A.P., India.
mastanmohd@gmail.com
[3]Research Scholar, Dept.of CSE, Dravidian University, Kuppam, A.P., India.
umar332@gmail.com

***Abstract* --- A wireless Ad-hoc network consists of wireless nodes communicating without the need for a centralized administration, in which all nodes potentially contribute to the routing process. In this paper, we report Fluctuations in channel quality effect the QoS metrics on each link and the whole end-to-end route. The interference from non-neighboring nodes affects the link quality. QoS is an essential component of ad-hoc networks. The most commonly studied QoS metrics are throughput, bandwidth, delay and jitter. Bandwidth is the QoS metric that has received the most attention in the QoS literature. The QoS requirements are typically met by soft assurances rather than hard guarantees from the network. Most mechanisms are designed for providing relative assurances rather than absolute assurances.**

Keywords: QoS, Ad-hoc, Throughput, Bandwidth, Delay, Jitter, 802.11.

I. INTRODUCTION

Wireless Ad-hoc network consists of wireless nodes communicating without the need for a centralized administration. The idea of such networking is to support robust and efficient operation ad-hoc wireless networks in which all nodes potentially contribute to the routing process, the fluctuations in channel quality effect the QoS metrics on each link and the whole end-to-end route. In ad-hoc networks, Quality of Service support is becoming an inherent necessity rather than an "additional feature" of the network. Wireless channel fluctuates rapidly and the fluctuations severely effect multi-hop flows. As opposed to the wired network, the capacity of the wireless channel fluctuates rapidly due to various physical layer phenomena including fading and multi-path interference. In addition, background noise and interference from nearby nodes further effect the channel quality. In ad-hoc networks, the end-to-end quality of a connection may vary rapidly as change in channel quality on any link may effect the end-to-end QoS metrics of multi-hop paths. The Packets contend for the shared media of the same stream at different nodes impacts the QoS metrics of a connection. Such contention arises as the wireless channel is shared by nodes in the vicinity. Interference effects are

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

pronounced in ad-hoc networks where typically a single frequency is used for communication in the shared channel. In Single hop infrastructured wireless networks frequency planning is mostly used where nearby base stations can be configured to function at different frequencies for reducing interference. Transmissions in the wireless media are not received correctly beyond the transmission range. But even beyond the transmission range, the remaining power may be enough to interfere with other transmission. So, interference from nonneighboring nodes may result in packet drops. In order to support QoS on multi-hop paths, QoS must be designed for the end-to-end path as well as for each hop. The physical and MAC layers are responsible for QoS properties on a single-hop. The routing layer is responsible for QoS metrics on an end-to-end route.

II.OVERVIEW OF IEEE 802.11 PHYSICAL LAYER

One of the fundamental challenges in wireless networks is the continuously changing physical layer properties of the channel. The physical layer of 802.11b can support multiple data rates. Depending on the channel quality the data rate can be altered to keep the bit error rate acceptable, as high data rates are also prone to high bit error rates.

The 802.11b standard operates in the 2.4 GHz band and supports 1, 2, 5.5 and 11 Mbps. For efficient use of a multi-rate physical layer, there have been several algorithms proposed at the physical layer. One of the algorithm which is closely tied to the MAC layer is Opportunistic Auto Rate (OAR) for improving throughput in the presence of multi-rate links in ad-hoc networks. The key idea is to send multiple packets when the channel rate is higher.

III.IMPORTANCE OF MEDIUM ACCESS LAYER

The original IEEE 802.11 [1] standard specifies the physical layer and the medium access layer mechanisms and provides a data rate up to 2 Mbps. Further the standards IEEE 802.11b modifies the physical layer part of the standard and increases the maximum data rates to 11 Mbps and 54 Mbps respectively. In this paper we discuss the basic 802.11 MAC layer functionality called Distributed Coordination Function (DCF) for distributed access to the shared medium. DCF is a natural choice for ad-hoc networks, as there is no centralized controller such as an

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

access-point. However, PCF can support QoS metrics in single-hop wireless networks due to its centralized design. Both DCF and PCF are enhanced in the upcoming standard 802.11e [2] that are designed for supporting QoS in WLANs.

IV. 802.11 DISTRIBUTED COORDINATION FUNCTION (DCF)

The DCF protocol attempts to provide equal access (in terms of number of packets) to all backlogged nodes that share a channel. In an ad-hoc network the throughput that a node obtains using DCF is a function of the number of neighbors that it has and the state of their queues (backlogged or not).

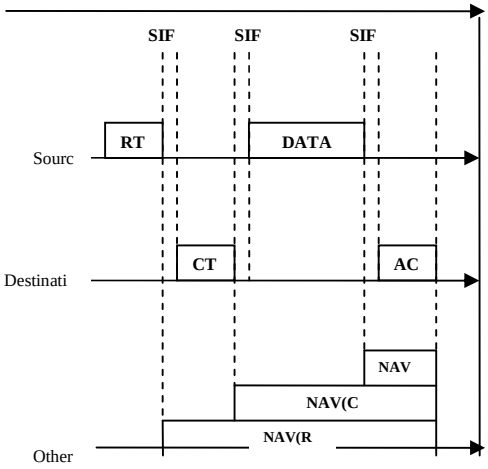


Figure 1: IEEE 802.11 DCF

Each node that has a packet to send picks a random slot for transmission in [0, cw], where cw is the contention window used for backoffs. Initially cw is set to cw_{min} . In the chosen slot the node sends a MAC layer control packet called RTS (request-to-send), to the receiver. If the receiver correctly receives the RTS and is not deferring transmission, it responds with CTS (clear-to-send). This is followed by transmission of the data packet by the sender, and a subsequent acknowledgment from the receiver. The transmissions of these four packets are separated by short durations called SIFS (Short Inter-Frame Space). The SIFS allows time for switching the transceiver between sending and receiving modes. The sequence of transmission of these four packets. The MAC header of all these packets contains a “duration” field indicating the remaining time till the end of the reception of the ACK packet. Based on this advertisement, the neighboring nodes update a data structure called NAV (Network Allocation Vector). This structure maintains the remaining time for which the node has to defer all transmissions.

If the packet transmission fails, the sender doubles its contention window (cw ←[2*cw-1]) and backs off before attempting a retransmission. The number of retransmissions is limited to

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

4 for small packets (including RTS packets) and 7 for larger (typically DATA) packets. If these counts are exceeded, the data packet is dropped and cw is reset to cw_{min} if the data packet is successfully delivered, both the sender and the receiver reset cw to cw_{min}.

V. PROPOSED QOS SUPPORT USING DCF BASED SERVICE DIFFERENTIATION

As it is difficult to provide absolute QoS guarantees, relative QoS assurance can be provided by service differentiation. However, to provide differentiated services, the 802.11 protocol needs to be modified. [3] proposes three ways to modify the DCF functionality of 802.11 to support service differentiation. The parameters that need to be modified to achieve service differentiation are.

1. *Backoff increase function:* Upon an unsuccessful attempt to send an RTS or a data packet, the maximum backoff time is doubled. More specifically the backoff time is calculated as follows:

$$Backoff_{time} = [\,2^{(2+i)} \times rand(\,)\,] \times Slot_{time}$$

Where i is the number of consecutive backoffs experienced for the packet to transmitted. To support different priorities, the backoff computation can be changed as follows:

$$Backoff_{time} = [\,P_j^{(2+i)} \times rand(\,)\,] \times Slot_{time}$$

where pj is the priority of node j

2. *DIFS:* As shown in Fig.1, this is the minimum interval of time required before initiating a new packet transmission after the channel has been busy. To lower the priority of a flow we can increase the DIFS (Distributed Coordination Function Inter Frame Spacing) period for packets of that flow. However, it is difficult to find an exact relation between the DIFS period for a flow and its throughput. Fig.2 shows the different DIFS values and the corresponding relative priorities.

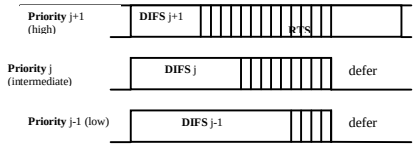


Figure 2: Service Differentiation using different DIFS values

3. *Maximum Frame Length:* Channel contention using the DCF functionality is typically used to send a single frame. By using longer frames, higher throughput can be provided to high-priority flows.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

VI. CONCLUSION

In this paper, the QoS issues discussed at various networking layers for ad-hoc networks. The physical layer and the MAC layers are primarily responsible for QoS metrics on each link and the whole end-to-end route. The DCF functionality of 802.11 is being extended and specifically designed for QoS support in multi-hop networks. The algorithm which is needed to be adapted for use in multi-hop ad-hoc networks is Opportunistic Auto Rate (OAR) for improving throughput in the presence of multi-rate links in ad-hoc networks.

QoS is currently an active research area in ad-hoc networks. However, there are several avenues that require further exploration for designing a QoS enabled ad-hoc network. For packets that traverse multiple hops, the end-to-end QoS is a function of the QoS metrics at each intermediate link. End-to-end QoS properties can be improved by designing a MAC layer that coordinates with other intermediate nodes on a multi-hop path. We find that QoS is an inherent component of ad-hoc networking and that there are several unsolved challenges that need to be addressed to design QoS enabled ad-hoc networks in future.

REFERENCES

[1] IEEE Std. 802.11. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, 1999.

[2] S. Mangold, S. Choi, G. R. Hiertz, O. Klein and B. Walke. Analysis of IEEE 802.11e for QoS Support in Wireless LANs .*IEEE Wireless Communications Magazine, Special Issue on Evolution of Wireless LANs and PANs*, Jul. 2003.

[3] I. Aad and C. Caselluccia. Differentiation mechanisms for IEEE 802.11. In *Proc. IEEE Infocom*, volume 2, pages 594–602, 1996.

[4] A.Veres, A. T. Campbell, M.Barry, and L. H. Sun. Supporting service differentiation in wireless packet networks using distributed control. *IEEE Journal on Selected Areas in Communications*, October 2001.

[5] B.Sadeghi,V.Kanodia,A.Sabharwal, and E. Knightly. Opportunistic Media Access for Multirate Ad-hoc Networks. In *Proc. ACM MOBICOM*, 2002.

AUTHORS PROFILE

Mohammed Ali Hussain received the Master's degree M.Sc Computer Science from Alagappa University in 2003. He received Master's degree M.Tech in Information Technology from Allahabad Deemed University in 2005. He received Ph.D. degree In Computer Science from Magadh University, Bihar, India in 2008. He is doing Post Doctoral degree in Computer Science & Engineering from Acharya Nagarjuna University, Guntur, Andhra Pradesh, India. He is currently an Associate Professor in the Department of Computer Science in Nimra College of Engineering & Technology, Vijayawada, Andhra Pradesh, India. He had published several papers in National and International Conferences & International Journals. His research interests are Wireless Networks with specialization in Quality of Service (QoS) in IEEE 802.11 Wireless LANs. & Ad-Hoc Networks. He is a member of IACSIT and ISTE.

Mohammed Mastan received the Master's degree in Computer Applications from Kakatiya University, Warangal in 2006. He received Master's degree in M.Tech Computer Science & Engineering from JNT University, Hyderabad in 2008. He is pursuing Ph.D. in Computer Science & Engineering from JNT University, Kakinada, Andhra Pradesh, India. He is currently as Asst.Professor in Department of Computer Science & Engineering in Nimra College of Engineering & Technology, Vijayawada, Andhra Pradesh, India. He has published several papers in National and International Conferences. His research interests are Computer Networks & Wireless Networks.

Syed Umar received the B.Tech degree Electronics and Communication Engineering from JNT University, Hyderabad in 2003. He received Master's degree M.Tech in Computer Science & Engineering from JNT University, Hyderabad in 2008. He is pursuing Ph.D. in Computer Science from Dravidian University, Kuppam, Chittoor Dist, Andhra Pradesh, India. He is currently an Asst.Professor in Department of Computer Science & Engineering in Nimra College of Engineering & Technology, Vijayawada, Andhra Pradesh, India. He had published several papers in National and International Conferences. His research interests are Computer Networks & Wireless Networks.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

A New Collaborative Web Recommendation Systems based on Association Rule Mining

A. Kumar
Research Scholar, Department of Computer Science & Engineering , Sathyabama University, Chennai, India
email:akr2020av@yahoo.in

Dr. P. Thambidurai
Principal, Perunthalaivar Kamarajar Institute of Engineering & Technology, Karaikal, India.
email: ptdurai58@yahoo.com

***Abstract*— Massive development of internet in recent years necessitate the development of recommender systems that turn out to be user friendly in web applications. Recommender systems make an effort to outline user preferences over items, and model the relation between users and items. There are two elemental approaches that can be applied when generating recommendations systems. They are content based web recommender system and the other is collaborative web recommender system. This proposed paper presents a method of developing a collaborative web recommendation systems using association rule mining. The association rules were applied to personalization based on web usage data. The method utilize apriori algorithm to generate association rules. In general association rule mining is a technique common in data mining that attempts to discover patterns of products that are purchased together. The greater part of web page recommender systems that were proposed earlier utilized collaborative filtering. Web Content Recommendation has been an active application area for Information Filtering, Web Mining and Machine Learning research. The future work explains some of the modifications using other algorithms to generate the association rules that can be adopted on existing web recommendation system to make them functionally more effective. In order to explore the performance of the proposed web recommendation system experiments were conducted on available dataset. The performance of the proposed approach is best illustrated by comparing it with K-nearest neighboring algorithm.**

***Keywords*---Association Rules, Apriori Algorithm, Collaborative Recommender System, Information Filtering, Machine Learning, and Web Mining.**

I. INTRODUCTION

The extent of the Internet is getting larger and larger in modern years. Therefore it is obligatory that a user need to expend much time to select indispensable information from large amount of web pages created every day. Addressing this problem, several web page recommender systems are constructed which automatically selects and recommends web pages suitable for user's support. The majority of web page recommender systems that was proposed earlier utilized collaborative filtering [1], [2], and [3]. Collaborative filtering is often used in general product recommender systems, and consists of the following stages. The foremost stage in collaborative filtering is to analyze users purchase histories in order to extract user groups which have similar purchase patterns. Then recommend the products that are commonly preferred in the user's group [4].

In general the Recommender Systems (RS) uses the opinion of members of a community to facilitate individuals identify the information most likely to be interesting to them or pertinent to their needs. This can be achieved by drawing on user preferences and filtering the set of feasible options to a more manageable subset. Every Web Recommendation Systems have its own advantages and limitations [5]. Moreover the assignment of recommender systems is to recommend items that fit a user's taste, in order to help the user in selecting/purchasing items from a devastating set of choices. Such systems have immense importance in applications such as e-commerce, subscription based services, information filtering, web services etc.

There are two fundamental approaches that can be applied when generating recommendations. Content based approaches profile users and items by identifying their characteristic features, such us demographic data for user profiling, and product information/descriptions for item profiling. The profiles are used by algorithms to unite user interests and item descriptions when generating recommendations [6]. Web Content Recommendation has been an active application area for Information Filtering, Web Mining and Machine Learning research. This proposed paper presents a method of developing a collaborative web recommendation systems using association rule mining. The method utilize apriori algorithm to generate association rules. It also explains some of the baseline algorithms that are used in developing the web recommendation systems. The future work explains some of the modifications using other algorithms to generate the association rules that can be adopted on existing web recommendation system to make them functionally more effective.

The remainder of this is organized as follows. Section 2 discusses various collaborative web recommendation systems that were earlier proposed in literature. Section 3 explains the proposed work of developing a web recommendation system

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

using association rules generated from apriori algorithm. Section 4 illustrates the results for experiments conducted on different dataset in evaluating the performance of the proposed web recommendation system. Section 5 concludes the paper with fewer discussions.

II. RELATED WORK

In general, the collaborative recommendation systems can be grouped into four categories. On the basis of its temporal and spatial characteristics, each system can be either synchronous or asynchronous, and either local or remote. Conversely, the most significant difference between these different collaborative Web recommendation systems is the method used to extort user preferences from personal information. This section of the paper discusses various methods proposed earlier in literature for a collaborative web recommendation system.

Chen et al. in [7] proposed a Gradual Adaption Model for a Web recommender system. The model is used to track users' center of attention and its transition by analyzing their information access behaviors, and recommend appropriate information. The web pages admittance by users are classified by the concept classes, and grouped into three terms of short, medium and long periods, and two categories of significant and incomparable for each concept class, which are used to describe users' focus of interests, and to institute reprocess probability of each concept class in each term for each user by Full Bayesian Estimation as well. According to the reuse probability and period, the information that a user is likely to be interested in is recommended. They proposed a new approach by which short and medium periods are determined based on dynamic sampling of user information access behaviors.

Niwa et al. in [8] described a web page recommender system based on Folksonomy mining. They projected a way to assemble a new type of web page recommender system covering all over the Internet, by using Folksonomy and Social Bookmark which are getting very well-liked in these days. a new way to express users' preference to web pages was formulated by mining tag data of Folksonomy. Folksonomy is a new classification technique which may take place of past taxonomy. Social Bookmark (SBM) is a variety of web services on which users can divide up their bookmarks. Anyone can see anyone's bookmark on SBM. In order to solve some problems faced by conventional recommender systems, they expressed users' web page preference by "affinity level between each user and each tag." By this approach, users' preferences are abstracted and it becomes easier to find similar users. Clustering can also solve the problem of "tag redundancy in Folksonomy."

A hybrid web recommender system was described by Taghipour et al. in [9]. They exploit an idea of combining the conceptual and usage information to enhance a reinforcement learning framework, primarily devised for web recommendations based on web usage data. Moreover the combination can improve the quality of web recommendations. A hybrid web recommendation method is proposed by making use of the conceptual relationships among web resources to derive a novel model of the problem, enriched with semantic knowledge about the usage behavior. With their proposed hybrid model for the web page recommendation problem they revealed the pertinent and flexibility of the reinforcement learning framework in the web recommendation domain, and demonstrated how it can be extended in order to incorporate various sources of information. Their test results suggested that the method can improve the overall quality of web recommendations.

An intelligent recommender system was projected by kavitha devi et al. in [10]. They designed and implemented an Intelligent Collaborative Recommender System (ICRS) to map users' needs to the items that can persuade them. A methodology is used to animatedly modernize the accuracy factor based on user intelligence. The diverse approaches for recommendation are categorized as memory-based and model-based approaches. Memory-based systems suffer from data sparsity and scalability problems, whereas model-based approaches are liable to bind the range of users. Therefore they integrated these approaches to overcome their limitations. They applied the collaborative filtering approach for recommendations. Recommendations are made more accurate by applying regression to weighted aggregated predictions. Mean Absolute Error Metrics was considered for evaluating the performance of their proposed system. This approach thus alleviates scalability and sparsity problems and offers accurate recommendations.

Cheng et al. in [11] developed a two stage collaborative recommender system. They proposed a chronological pattern based collaborative recommender system that predicts the customer's time-variant acquisition behavior in an e-commerce environment where the customer's purchase patterns may change gradually. A new two-stage recommendation process is developed to envisage customer behavior for the selection of different categories, as well as for product items. Their study is the first to recommend time-decaying sequential patterns within a collaborative recommender system. Their experimental results revealed that the proposed system outperforms the traditional collaborative system.

Lin et al. in [12] described an efficient Adaptive-Support Association Rule Mining for Recommender Systems. They investigated the utilization of association rule mining as an underlying technology for collaborative recommender systems. Association rules have been used with sensation in other domains. Nevertheless, most currently existing association rule mining algorithms were designed with market basket analysis in mind. They described a collaborative recommendation technique based on a novel algorithm distinctively designed to excavate association rules for this rationale. The main advantage of their proposed approach is that their algorithm does not require the minimum support to be specified in advance. Rather, a target range is given for the number of rules, and the algorithm adjusts the minimum support for each user in order to obtain a rule set whose size is in the desired range. Moreover they employed associations between users as well as associations between items in making recommendations. The experimental evaluation of a

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

system based on their algorithm revealed that its performance is significantly better than that of traditional correlation-based approaches.

Jung in [13] described a user-support method based on the distribution of knowledge with other users through the collaborative Web browsing, focusing exclusively on the user's interests extracted from users' own bookmarks. More prominently, they focused on those items of information which are associated to the user's interests. In collaborative Web browsing, they considered that recognizing the user's interests is an extremely essential mission. Furthermore, asking applicable information for other users, filtering the query results, and recommending them are additional most important tasks that have to be unreservedly conducted by them in [13]. Based on the personalized TF-IDF proposal they introduced the extended application of a BIS Agent, which is a bookmark sharing agent system. Moreover they implemented an ontological supervisor which can perform the semantic analysis of the Web sites pointed to by these bookmarks. They also designed a multi-agent system that consists of a facilitator agent and many personal agents. The main limitation of this system is that it does not consider the privacy problems related with sharing personal information of the user.

A novel recommender system was formulated by Marko et al. in [14]. Their approach named, "Fab" is a recommendation system designed to help users sift through the mammoth amount of information obtainable in the World Wide Web. Their proposed approach is the combination of content-based filtering and collaborative filtering methods. The combination exploits the advantages of both the methods thereby avoiding the shortcomings. Fab's hybrid structure allows for automatic recognition of emergent issues relevant to various groups of users. It also enables two scaling problems, pertaining to the rising number of users and documents, to be addressed. In general the content-based approach to recommendation has its pedigree in the information retrieval (IR) community, and utilizes many of the same techniques. The collaborative approach computed the similarity of the users rather than computing the similarity of the items. They maintained user profiles content analysis and directly compared these profiles to determine similar users for collaborative recommendation. The process of recommendation can be partitioned into two stages: collection of items to form a manageable database or index, and subsequently selection of items from this database for particular users. The experimental results using the hybrid Fab system achieved higher accuracy.

III. PROPOSED APPROACH

A. Association Rule Mining

In general association rule mining is a technique common in data mining that attempts to discover patterns of products that are purchased together. The proposed approach adapts the Apriori algorithm [15] to collaborative filtering in an attempt to discover patterns of items that have common ratings. Association rules capture relationships among items based on patterns of co-occurrence across transactions. The association rules were applied to personalization based on web usage data in [16]. Considering each user profile as a transaction, it is possible to use the Apriori algorithm [15] and generate association rules for groups of commonly liked items.

Given a set of user profiles U and a set of item sets $I = \{I_1, I_2 \ldots I_k\}$, the support of an item set $I_i \in I$ is defined as $\sigma(I_i) = |\{u \in U : I_i \subseteq u\}| / |U|$. Item sets that satisfy a minimum support threshold are used to generate association rules. These groups of items are referred to as frequent item sets. In addition, association rules that do not satisfy a minimum lift threshold are shortened. If there is not adequate support for a particular item, that item will never come into view in any recurrent item set. The suggestion is that such an item will never be recommended. The subject of coverage is a tradeoff. Lowering the support threshold will make sure that more items can be recommended, but at the hazard of recommending an item without enough evidence of a pattern.

Ahead of performing association rule mining on a collaborative filtering dataset, it is indispensable to discretize the rating values of each user profile. Therefore first subtract each user's average rating from the ratings in their profile to attain a zero-mean profile. Next, give a discrete category of "like" or "dislike" to each rated item in the profile if it's rating value is $>$ or $\leq$ zero, respectively. As a result of discretizing the dataset the total number of features used in the analysis is doubled. It is clear that a collaborative recommender must take such preference into description or risk recommending an item that is rated often, but disliked by consensus. Another possibility is that one association rule may add item 'i' to the candidate set with a "like" label, while another rule may add the identical item with a "dislike" label. There is not an ultimate solution in this case, but we have chosen to presuppose that there are opposing forces for the recommendation of the item. This implementation subtracts the confidence value of the "dislike" label from the confidence value of the "like" label. The search for item sets is facilitated by storing the frequent item sets in a directed acyclic graph, called a Frequent Item set Graph [16].

Given a objective user profile 'u', it is necessary to execute a depth-first search of the Frequent Item-set Graph. When we arrive at a node whose repeated item set I_n is not enclosed in 'u', the item $i \in I_n$ not found in 'u' is added to the candidate set 'C' and search at the current branch is finished. Note that the item set of the parent node I_p to I_n must be restricted in 'u' by definition, and because I_n is of size d + 1 where I_p is size 'd', there can be only one item $i \in I_n$ that is not contained in 'u'. It follows that $I_n = I_p \cup \{i\}$ and the two nodes correspond to the rule $I_p \Rightarrow \{i\}$. Then calculate the confidence of the rule as $\sigma(I_n)/\sigma(I_p)$. The candidate $i \in C$ is stored in a hash table along with its confidence value. If it already exists in the hash table, then the highest confidence value takes precedent.

After finishing point of the depth-first search, all promising candidates for the target user 'u' are enclosed in 'C', including items labeled "dislike". Consecutively to accurately characterize an estimated negative implication, items labeled "dislike" is given a recommendation score that is the negation of the confidence value. As a final step, the

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

candidate set 'C' is arranged according to the recommendation scores and the top N items are returned as a recommendation.

B. Baseline Algorithms

K Nearest Neighbour

The customary k-nearest neighbor algorithm is broadly used and reasonably precise [17]. Resemblance is calculated using Pearson's correlation coefficient, and the k most analogous users that have rated the target item are selected as the neighborhood. This make note that a target user may have a different neighborhood for each target item. It is also general to filter neighbors with similarity below a specified threshold. This prevents prophecy being based on very remote or negative correlations. After identifying a neighborhood, Resnick's algorithm is used to calculate the prediction for a target item 'I' and target user 'u'.

K Means Clustering

A standard model-based collaborative filtering algorithm uses k-means to cluster similar users. Given a set of user profiles, the space can be partitioned into k groups of users that are close to each other based on a measure of similarity. The discovered user clusters are then applied to the user based neighborhood formation task, rather than individual profiles. To make a recommendation for a target user 'u' and target item 'i' it is essential to select a neighborhood of user clusters that have a rating for 'I' and whose aggregate profile is most similar to 'u'. This neighborhood represents the set of user segments that the target user is most likely to be a member, based on a measure of similarity. Pearson's correlation coefficient is implemented effectively to perform this task.

Probabilistic Latent Semantic Analysis

In general the Probabilistic latent semantic analysis (PLSA) models [18] present a probabilistic approach for characterizing latent or hidden semantic associations among co-occurring objects. PLSA can be applied to the creation of user clusters based on web usage data. The proposed approach has adapted this technique to the context of collaborative filtering [19]. Moreover the Expectation-Maximization (EM) algorithm is used to perform maximum likelihood parameter estimation. In the expectation step, posterior probabilities are computed for latent variables based on current estimates. In the maximization step, Lagrange multipliers [20] are used to obtain the re-estimated parameters. Therefore iterating the expectation and maximization steps monotonically increases the total likelihood of the observed data L (U, I), until a local optimum is reached.

IV. Experimental Results

In order to evaluate the robustness of our recommendation algorithm based on association rule mining a data set is taken into account. This dataset consists of 100,000 ratings on 1682 movies by 943 users. All ratings are integer values between one and five, where one is the lowest (dis-liked) and five is the highest (liked). Initially the accuracy of the proposed association rule mining based recommender system is analyzed. To estimate the recommendations, 10-fold cross-validation on the entire dataset and without attack profiles was performed. Since Apriori selects recommendations from only among those item sets that have met the support threshold, it will by necessity have lower coverage than our baseline algorithms. There will be some items that do not appear in the Frequent Item set Graph, and about which the algorithm cannot make any predictions. This may the issues that arise in most of the baseline algorithms.

The Apriori algorithm would therefore lend itself best to scenarios in which a list of top recommended items is sought, rather than a rating prediction scenario in which the recommender must be able to estimate a rating for any given item. The selectivity of the algorithm may be one reason to expect it will be relatively robust - it will not make recommendations without evidence that meets the minimum support threshold. However, the performance of Apriori and PLSA are superior to k-means at large attack sizes. Robustness of the Apriori algorithm may be moderately due to lower coverage. However, this does not account for the flat trend of hit ratio with respect to attack size.

Only the Apriori algorithm holds steady at large filler sizes and is essentially unaffected. As with attack size, the reason that filler size does not affect the robustness of the algorithm is because adding more filler items does not change the probability that multiple attack profiles will have common item sets. The fact that a profile's ratings are discretized to categories of "like" and "dislike" means that an attack profile with 100% filler size will cover exactly half of the total features used in generating frequent item sets. Therefore, it is very unlikely that multiple attack profiles will result in mutual reinforcement. Apriori has also exhibited improved robustness compared to the other algorithms against certain attacks.

The Apriori algorithm appears to have the same robustness as the other model-based algorithms at small attack sizes. Although the performance of Apriori is not ideal against a segment attack, certain scenarios may minimize the performance degradation in practice. In particular, a recommender system with a very large number of users is somewhat buffered from attack. The algorithm is quite robust through a 5% attack, and is comparable to both k-means and PLSA. The robustness of Apriori is not drastically reduced until attack size is 10% or greater. Table 1 shows the results of normalized mean absolute error evaluated for proposed approach and K nearest neighbor. Similar table 2 represents the coverage comparison of the proposed approach using apriori algorithm and K nearest neighbor. Figure 1 (a) shows the comparison of the proposed approach and k-nn algorithm in terms of their mean absolute error. Figure 1 (b) represents the comparison of apriori and k-nn algorithm in terms of coverage. The results revealed that the proposed approach using association rule mining performed well in recommendation by determining the user' needs.

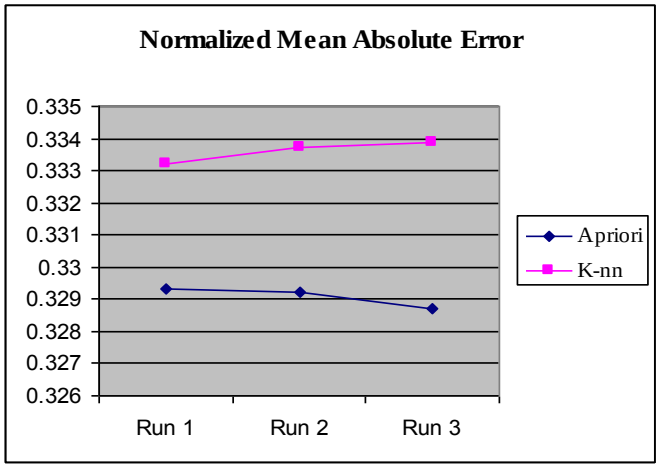
http://sites.google.com/site/ijcsis/
ISSN 1947-5500

TABLE.1. NORMALIZED MEAN ABSOLUTE ERROR

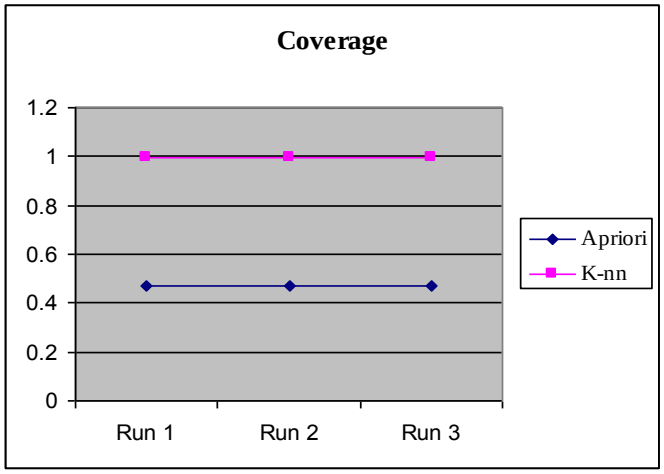
Approach	Run 1	Run 2	Run 3	Mean
Apriori	0.3293	0.3292	0.3287	0.3291
K-nn	0.3332	0.3337	0.3339	0.3336

TABLE.2. COVERAGE

Approach	Run 1	Run 2	Run 3	Mean
Apriori	0.4701	0.4716	0.4718	0.4712
K-nn	0.9942	0.9941	0.9942	0.9942



(a)



(b)

Figure.1 (a) & (b) represents comparison of Apriori and K-nn Algorithm in terms of Normalized Mean Absolute Error and Coverage respectively

V. CONCLUSION

The scope of the Internet is getting larger and larger in recent years. Therefore it is compulsory that a user need to disburse much time to decide on necessary information from large amount of web pages created every day. Addressing this problem, several web page recommender systems are constructed which automatically selects and recommends web pages appropriate for user's support. The greater part of web page recommender systems that were proposed earlier utilized collaborative filtering. This proposed paper presents a method of developing a collaborative web recommendation systems using association rule mining. In general association rule mining is a technique common in data mining that attempts to discover patterns of products that are purchased together. The association rules were applied to personalization based on web usage data. The method utilize apriori algorithm to generate association rules. The Apriori algorithm would therefore provide itself best to scenarios in which a list of top recommended items is required, rather than a rating prediction scenario in which the recommender must be able to approximate a rating for any given item. The results revealed that the proposed approach using association rule mining performed well in recommendation by determining the user' needs. Future work mainly concentrates on determining the mutual reinforcement between common item sets thereby ehnancing the accuracy of the recommender system.

REFERENCES

[1] J. Li, and O. Zaiane, "Combining Usage, Content, and Structure Data to Improve Web Site Recommendation," Proceedings of WebKDD-2004 workshop on Web Mining and Web Usage, In 5th International Conference on Electronic Commerce and Web Technologies, pp. 305-315, 2004.

[2] P. Kazienko, and M. Kiewra, "Integration of relational databases and Web site content for product and page recommendation," International Symposium on Database Engineering and Applications Symposium, IDEAS '04, pp. 111-116, 2004.

[3] N. Golovin, and E. Rahm, "Reinforcement Learning Architecture for Web Recommendations," Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC'04), vol. 2, p. 398, 2004.

[4] Jonathan L. Herlocker, Joseph A. Konstan, Loren G. Terveen, and John T. Riedl, "Evaluating collaborative filtering recommender systems," ACM Transactions on Information Systems, vol. 22, no. 1, pp. 5-53, 2004.

[5] Punam Bedi, Harmeet Kaur, and Sudeep Marwaha, "Trust based Recommender System for the Semantic Web," Proceedings of the 20th international joint conference on Artificial intelligence, pp. 2677-2682, 2007.

[6] Gabor Takacs, Istvan Pilaszy, Bottyan Nemeth, and Domonkos Tikk, "Scalable Collaborative Filtering Approaches for Large Recommender Systems," The Journal of Machine Learning Research, vol. 10, pp. 623-656, 2009.

[7] Jian Chen, Roman Y. Shtykh, and Qun Jin, "A Web Recommender System Based on Dynamic Sampling of User Information Access Behaviors," Ninth IEEE International Conference on Computer and Information Technology, vol. 2, pp. 172-177, 2009.

[8] Satoshi Niwa, Takuo Doi, and Shinichi Honiden, "Web Page Recommender System based on Folksonomy Mining," Proceedings of the Third International Conference on Information Technology: New Generations, IEEE Computer Society, pp. 388-393, 2006.

[9] Nima Taghipour, and Ahmad Kardan, "A hybrid web recommender system based on Q-learning," Proceedings of the 2008 ACM symposium on Applied computing, pp. 1164-1168, 2008.

[10] M.K. Kavitha Devi, and P. Venkatesh, "ICRS: an intelligent collaborative recommender system for electronic purchasing,"

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

International Journal of Business Excellence 2009, vol. 2, no. 2, pp. 179-193, 2009.
[11] Cheng-Lung Huang, and Wei-Liang Huang, "Handling sequential pattern decay: Developing a two-stage collaborative recommender system," Elsevier, Electronic Commerce Research and Applications, vol. 8, no. 3, pp. 117-129, 2009.
[12] Weiyang Lin, Sergio A. Alvarez, and Carolina Ruiz, "Efficient Adaptive-Support Association Rule Mining for Recommender Systems," Journal on Data Mining and Knowledge Discovery, vol. 6, no. 1, pp. 83-105, 2004.
[13] Jason J. Jung, "Collaborative Web Browsing Based on Semantic Extraction of User Interests with Bookmarks," Journal of Universal Computer Science, vol. 11, no. 2, pp. 213-228, 2005.
[14] Marko Balabanovic, and Yoav Shoham, "Fab: content-based, collaborative recommendation," Communications of the ACM, vol. 40, no. 3, pp. 66-72, 1997.
[15] R. Agrawal, and R. Srikant, "Fast algorithms for mining association rules," In Proceedings of the 20th International Conference on Very Large Data Bases (VLDB'94), Santiago, Chile, September 1994.
[16] M. Nakagawa, and B. Mobasher, "A hybrid web personalization model based on site connectivity," In Web KDD Workshop at the ACM SIGKKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, August 2003.
[17] J. Herlocker, J. Konstan, A. Borchers, and J. Riedl, "An algorithmic framework for performing collaborative filtering," In Proceedings of the 22nd ACM Conference on Research and Development in Information Retrieval (SIGIR'99), Berkeley, CA, August 1999.
[18] T. Hofmann, "Probabilistic latent semantic analysis," In Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence, Stockholm, Sweden, July 1999.
[19] B. Mobasher, R. Burke, and J. Sandvig, "Model-based collaborative filtering as a defense against profile injection attacks," In Proceedings of the 21st National Conference on Artificial Intelligence, pp. 1388–1393, AAAI, July 2006.
[20] T. Hofmann, "Unsupervised learning by probabilistic latent semantic analysis," Machine Learning Journal, vol. 42, no. 1, pp. 177–196, 2001.

Author Biography

A. Kumar has around 21 years of experience in Information Technology and its Applications with expertise in Data mining, Information and Knowledge Management and Web Technology. He has published a number of papers in peer-reviewed journals and conferences. He has actively served as a reviewer for many leading International Journals and conferences. He has delivered lectures in the areas of Web Recommendation Systems, Knowledge Management for Seminars and conferences. Currently serving as Head in Computer Science and Engineering, Perunthalaivar Kamarajar Institute of Engineering and Technology, India and Pursuing Ph.D., in the area of Web Recommendation Systems at Sathyabama University, Chennai.

Dr. P. Thambidurai is Principal, Perunthalaivar Kamarajar Institute of Engineering & Technology, Karaikal, India. He received the M.E (Computer Science)., from College of Engineering, Guindy, Chennai and Ph.D. in Computer Science from Alagappa University, Karaikudi, India in 1984 and 1995 respectively. He has published more than 100 papers in peer-reviewed journals and conferences. His research interests include Natural Language Processing and Data Compression. He has delivered lectures in the areas of NLP, Information Security and Image Compression in conferences. He organized and acted as a chair person for various seminars and conferences. He introduced many research programmes in Pondicherry University. He is a life member of Computer Society of India and Indian Society for Technical Education.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

SIMILARITY BASED IMPUTATION METHOD FOR TIME VARIANT DATA

F.Sagayaraj Francis, Vishnupriya.B, Vinolin Deborah Delphin, Saranya Kumari Potluri
Pondicherry Engineering College, Pondicherry, India

***Abstract*—The intent of any analysis is to make valid inferences regarding a population of interest. Missing data threatens this goal if it is missing in a manner which makes the sample different than the population from which it was drawn, that is, if the missing data creates a biased sample. Therefore, it is important to respond to a missing data problem in a manner which reflects the population of inference. This paper deals with the proposal of an efficient method of filling missing data called Similarity based Imputation Method (SIM). SIM processes the target segment by extracting its features and based on the extracted features the target segment is classified into its appropriate cluster which has complete data segments, which are similar to the target segment. Now from the similar segments within the identified cluster, the most identical segment is found using similarity measure and the values substituted from that complete segment**

Keywords-- Imputation, Time variant Multi-dimensional data, Clustering, Feature Extraction, Similarity measure, Segment Matching

I. INTRODUCTION

A time series may be defined as a collection of readings belonging to different time periods, of some economic variable or composite of variables. It is a set of observations of a variable usually at equal intervals of time where time may be yearly, monthly, weekly, daily, hourly or even minute data. Hourly temperature reading, daily sales, monthly production, Earth's magnetic field variations are examples of time series. Time series analysis comprises methods for analyzing time series data in order to extract meaningful statistics and other characteristics of the data.

The primary purpose of the analysis of time series is to discover and measure all types of variations which characterize a time series. For efficient analysis of data, complete datasets are required. There is a possibility to miss out several observations due to unexpected events such as equipment failure or unexpected disturbances. Moreover, occurrence of missing observations in datasets is an actual yet challenging issue confronted in machine learning and data mining. Missing values may generate bias and affect the quality of the supervised learning process or the performance of classification algorithms. However, most learning algorithms are not well adapted to some application domains due to the difficulty with missing values (for example, Web applications) as most existing algorithms are designed under the assumption that there are no missing values in datasets. That implies that a reliable method for dealing with those missing values is necessary. Generally, dealing with missing values means to find an approach that can fill them and maintain the original distribution of the data as closely as possible. There are many approaches to deal with missing values described in, for instance:

1. Ignore objects containing missing values;
2. Fill the missing value manually
3. Substitute the missing values by a global constant or the mean of the objects
4. Obtain the most probable value to fill in the missing values

The first approach usually loses too much useful information, whereas the second one is time consuming and expensive and hence infeasible in many applications. The third and fourth approaches assume that all missing values are with the same value, probably leading to considerable distortions in data distribution. The method of imputation, however, is a popular strategy. In comparison to other methods, it uses as much information as possible from the observed data to predict missing values.

The approach used in this paper fills the missing data by imputation using a similarity measure. The method involves finding a similar segment that matches the segment with missing data best using similarity measures of the segments.

II. EXISTING METHODOLOGIES

In general there can be two approaches to handle missing data. The first approach to find missing or lost data is by computation. These methods are statistical or numeric in nature, such as Line/Parabola of best fit or Interpolation/Extrapolation. The methods heavily depend on the formation of a function/model/system to fill missing data. The formation of function/model/system is inherently complex and requires voluminous computations.

The second approach to find missing or lost data is by imputation. Imputation is the substitution of missing values in an incomplete dataset (target dataset) with values from a similar but complete dataset (source dataset). This approach uses similarity/distance/entropy measures to solve the problem. They can also be extended to predict patterns that are likely to occur. Traditional missing value imputation techniques can be roughly classified into two types.

- Parametric imputation (e.g., the linear regression) and
- Non-parametric imputation (e.g., non-parametric kernel-based regression method, Nearest Neighbor method).

The parametric regression imputation is superior if a dataset can be adequately modeled parametrically, or if users

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

can correctly specify the parametric forms for the dataset. For instance, the linear regression methods usually can treat well the continuous target attribute, which is a linear combination of the conditional attributes. However, when the actual relation between the conditional attributes and the target attribute is not known, the performance of the linear regression for imputing missing values is very poor. In real application, if the model is incorrectly specified, the estimations of parametric method may be highly biased and the optimal control factor settings may be miscalculated. Non-parametric imputation algorithms that can provide superior fit by capturing structure in the dataset if the actual distribution of a dataset is known. Using a non-parametric algorithm is beneficial when the form of relationship between the conditional attributes and the target attribute is not known a-priori.

While nonparametric imputation method is of low-efficiency, the popular nearest neighbor (NN) methods face two issues: (i) Each instance with missing values requires the calculation of the distances from it to all other instances in a dataset; and (ii) There are only a few random chances for selecting the nearest neighbor.

A. Feature Extraction

Time variant datasets are voluminous and hence handling raw time variant data is not viable. A solution to this would be to extract features from time series segments and deal with features rather than the raw data as such.

Generalized feature extraction relies on a collection of feature extractors that function independently of domain and application. For time series data, such feature extractors must be able to identify generally useful structures that emerge from the relationships between consecutive measurement values over time.

For traditional tabular data the similarity is often measured by attribute value similarity or even attribute-value equality [1]. For more complex data, e.g., geo- time series data, such simple similarity measures do not perform very well. So similarity of time series data should be based on certain characteristics of the data rather than on the raw data itself. Feature Extractor is characterized by its ability to capture fundamental trends and relationships, generate accurate approximations, represent the extracted structures compactly, support subsequent classification, and being domain independent [2]. But this technique only deals with finding the nature of a particular segment and does not extract the uniqueness of each segment and it does not maintain originality of the target segment. Therefore a more general time series feature extraction technique is to be adopted which would represent any time series.

A well established feature extraction technique using Discrete Fourier Transformation (DFT) for time series use only the first k coefficients, discarding the rest. This corresponds to saving only a rough sketch of the time series, because the first coefficients represent the overall nature of the segment [6,8]. If N is the number of observations, in single dimension,

$$X_k = \sum_{N=0}^{N-1} x_e \, e^{-\frac{2\Pi i}{N} kn} \qquad \text{where } k=0,\ldots,N\text{-}1 \quad (1)$$

Since the data set would be a bunch of time segments, the DFT would thus transform them from the time domain into frequency domain. It is easy and fast to compute. It preserves the distance between two objects. DFT ensures completeness of feature extraction [7].

B. Similarity Search

Given a set of time series segments, there are two types of similarity searches: (i) sub segment matching finds all the data sub segments of a larger segment that are similar to the given smaller segment and (ii) whole matching finds those segments that are similar to one another.

Two key aspects for achieving effectiveness and efficiency, when managing time series data are representation methods and similarity measures. Time series are essentially high dimensional data and dealing with raw data directly is very expensive in terms of processing and storage cost. Unlike canonical data types like ordinal or nominal variables, where the distance definition is straightforward, the distance between time series needs to be carefully defined in order to reflect the underlying similarity of such data. This is particularly desirable for similarity-based retrieval, classification and clustering of time series.

Many similarity measures are efficient in accurately quantifying the similarity between any two distinct objects. Some of the efficient similarity measures to state can be Bray-Curtis distance, Canberra Distance, Cosine Distance, Correlation Distance and Chessboard Distance. However these measures were not able to capture the similarity perfectly between any two time variant segments. These methods were lacking in way that they compare only the corresponding points of the two data sets, but a time varying dataset is unique in its nature of being recorded at discrete and equal intervals of time and hence similarity measures exclusively for time variant data are to be considered.

Past research, on the choice of distance/similarity function to obtain similarity measure between two segments that are time variant, reveal that it can be divided into two classes. The first and straightforward method is the Euclidean distance and its variants like Manhattan distance. The complexities of these measures are linear, they are easy to implement and parameter free.

However, since the mapping between the points of two time series is fixed in the above two methods, these distance measures are very sensitive to noise and misalignments in time, and are unable to handle local time shifting i.e., similar segments that are out of phase,

The second class includes the Dynamic Time Warping (DTW) [3]. Continuity is less important in DTW than in other pattern matching algorithms. DTW is an algorithm particularly suited to matching segments with missing information, provided there are long enough segments for

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

matching to occur. Optimal alignment [4,10] i.e., minimum distance time warp path is obtained by allowing assignment of multiple successive values of one time series to a single value of the other time series and therefore it can be calculated on time series of different lengths.

Dynamic time warping is an algorithm for measuring similarity between two segments which may vary in time. DTW aligns two time series in the way some distance measure is minimized as shown in Fig. 1. It can efficiently handle noise and misalignments in time, and are able to handle local time shifting i.e., similar segments that are out of phase [9].

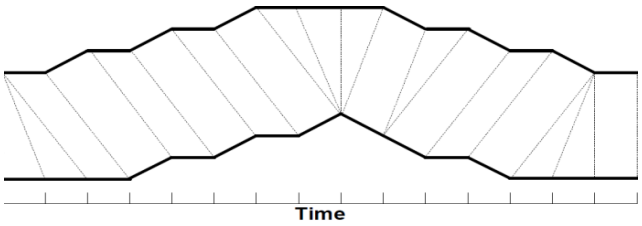


Figure. 1 A warping between two time series

In Fig. 1, each vertical line connects a point in one time series to its correspondingly similar point in the other time series. The lines actually have similar values on the y-axis but have been separated so the vertical lines between them can be viewed more easily. If both of the time series in Fig. 1 were identical, all of the lines would be straight vertical lines because no warping would be necessary to 'line up' the two time series. The warp path distance is a measure of the difference between the two time series after they have been warped together, which is measured by the sum of the distances between each pair of points connected by the vertical lines in Fig.1. Thus, two time series that are identical except for localized stretching of the time axis will have DTW distances of zero. Given two time series X, and Y, of lengths |X| and |Y|; X= $x_1,x_2,\ldots,x_i,\ldots,x_{|x|}$ and Y= $y_1,y_2,\ldots,y_i,\ldots,y_{|y|}$, a warp path W is given by

W= $w_1,w_2,\ldots,w_i,\ldots,w_k$, where $\max(|X|,|Y|) \le k < |X|+|Y|$

k is the length of the warp path and the k^{th} element of the warp path is $W_k=(i,j),w_{k+1}=(i_k,j_k)$ and *i* is an index from time series X, and *j* is an index from time series Y. The warp path must start at the beginning of each time series at $w_1 = (1, 1)$ and finish at the end of both time series at $w_K = (|X|, |Y|)$. This ensures that every index of both time series is used in the warp path. There is also a constraint on the warp path that forces *i* and *j* to be monotonically increasing in the warp path, which is why the lines representing the warp path in Fig. 1 do not overlap. Every index of each time series must be used. More formally, the optimal warp path is the warp path is the minimum-distance warp path, where the distance of a warp path W is

$$\text{Dist(w)} = \sum_{k=1}^{k=K} Dist\,(w_{ki} w_{kj}\,,) \qquad (2)$$

Dist(W) is the distance, of warp path W, and *Dist(w_{ki}, w_{kj})* is the distance between the two data point indexes in the k^{th} element of the warp path.

Thus DTW gives a quantified measure of similarity between any two instances and smaller the value more is the similarity between them. So to impute any segment with missing data, it is required to compute DTW between the target segment and every other complete segment in the dataset. But since the dataset is very large it is cumbersome to calculate similarity measure with every other segment. Thus there arises a need for clustering wherein the raw data and their extracted features can be grouped as clusters each of which contains similar instances.

C. Clustering

Clustering is a division of data into groups of similar objects. Each group or a cluster consists of objects that are similar between themselves and dissimilar to objects of other groups. When the dataset under consideration is very large, manual handling of data proved to be inefficient. So clustering techniques are adopted to handle the data and thereby minimize the computations.

The common clustering techniques include hierarchical clustering, partitioning clustering, grid based clustering and constraint based clustering. Some of the hierarchical clustering techniques are agglomerative clustering and divisive clustering. k-medoids, k-means, probabilistic and density based clustering techniques fall under the category of partitioned clustering.

III. SIMILARITY BASED IMPUTATION METHOD

The key objective of this paper is to fill out missing data that may be lost due to some inevitable reasons. Every data point adds to the machine learning or data mining process which implies that no data point can be ignored. So the need to fill missing data by some efficient method is needed that involves less computation and would fill in the missing observations more logically and accurately. This section discusses a new method to fill missing data called **Similarity based Imputation Method (SIM)** which involves filling by similarity search. The idea is to fill missing data by finding similar complete segments for the segments with missing data and impute them.

SIM takes as input a large time variant dataset with missing segments. The imputation process consists of two phases. The first phase collects the significant parameters of the dataset and the second uses these values to impute. The steps in the first phase are:

1. Divide the entire dataset into segments of length *l*.
2. Except for the segments with missing data, compute segment parameters for each of the segment in the dataset.
3. Cluster the segments based on their segment parameters.

The second phase that imputes the missing values consists of the following steps. The steps are applied on each segment with missing data.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

1. Compute segment parameters and classify the segment into an appropriate cluster.
2. Divide each segment into sub segments.
3. Extract DFT coefficients of all the sub segments.
4. Using DFT coefficients find the most similar segment to the segment with missing data. Impute values from the complete segment into the segment with missing data

A lot of segment parameters may be used based on which the segments can be clustered. A few possible parameters may be count of local maxima, count of local minima, position of global maxima, position of global minima, global maxima value and global minima value. For example for the sample dataset (223.9, 223.7, 223.6, 223.4, 223.3, 223.3, 223.2, 223.1, 223.2, 223.4, 223.5, 223.6, 223.8, 223.7, 223.8, 223.8, 223.8, 223.7, 223.7, 223.9, 223.9, 224.0, 224.1, 224.2), the above said parameters are 7, 6, 24, 8, 224.2, 223.1, respectively.

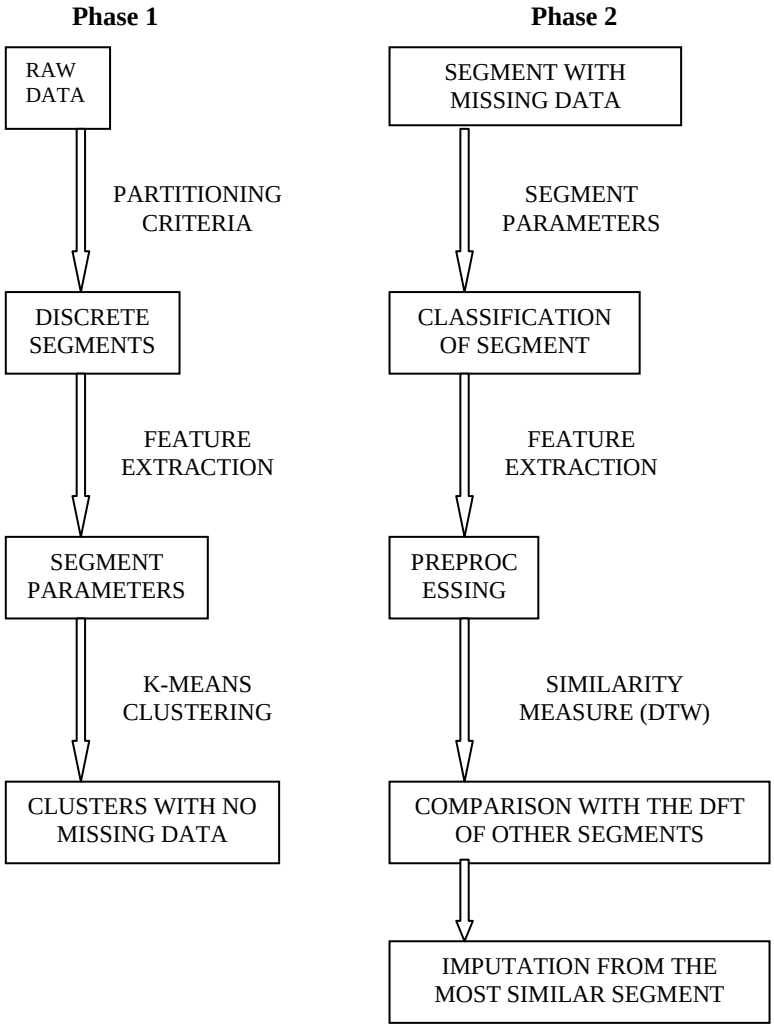


Figure. 2 Phase 1 and Phase 2 of imputation process

To compute the DFT coefficients, each segment is divided into three sub segments. Let *a* and *b* be the positions from and to in the incomplete segment where the data are missing. The data from the first to $(a-1)^{th}$ position forms the first sub segment; data from a^{th} to b^{th} position forms the second sub segment and the rest form the third sub segment of the segments. If (223.9, 223.7, 223.6, 223.4, 223.3, 223.2, -1, -1, -1, -1, -1, -1, 223.7, 223.8, 223.8, 223.8, 223.8, 223.7, 223.7, 223.9) represents the segment and -1 represents the missing value, then *a* and *b* of all the segments would be 7 and 12. The DFT coefficients of the upper and lower sub segments of the dataset are (591.283, -0.041,-0.091, 0.097, 0.312, 0.663) and (592.044, -0.252, -0.25, -0.178,0.039, 0.507)

The identification of the similarity of the sub segments involves the computation of similarity measure DTW between every pair of sub segments. The candidate for the imputation is the segment that is most similar to upper and lower sub segments of the incomplete segment. The Fig. 2 explains the steps pictorially.

IV. EXPERIMENTAL EVALUATION

The dataset considered for experimental evaluation of the method proposed in this paper is one of the three measures of earth's magnetic field variations which are measured along three axes as H- the horizontal intensity, Z- the vertical intensity and D the dip. The frequency of observation is one minute. The parameter that was used for evaluating the result is *Relative Error Percentage* (REP). If AV is the actual value and IV is the imputed value REP is given by

$$REP = \frac{\text{AV} - \text{IV}}{AV} \quad (3)$$

Table 1 Relative Error Percentage of various experimental setups

% of Missing Points	Region of Missing Points	Number of Segments						
		100	200	500	1000	5000	10000	20000
5%	Upper	0.42	0.17	0.17	0.17	0.17	0.17	0.17
	Middle	0.25	0.128	0.128	0.128	0.128	0.128	0.085
	Lower	0.553	0.043	0.043	0.043	0.043	0.043	0.043
4%	Upper	0.128	0.06	0.06	0.06	0.06	0.06	0.06
	Middle	0.17	0.128	0.128	0.128	0.128	0.128	0.128
	Lower	0.46	0.045	0.045	0.045	0.045	0.045	0.045
3%	Upper	0.06	0.06	0.06	0.06	0.06	0.06	0.06
	Middle	0.17	0.128	0.128	0.128	0.128	0.128	0.085
	Lower	0.468	0.043	0.043	0.043	0.043	0.043	0.043
2%	Upper	0.426	0.06	0.06	0.06	0.06	0.06	0.06
	Middle	0.17	0.128	0.128	0.128	0.128	0.128	0.085
	Lower	0.468	0.043	0.043	0.043	0.043	0.043	0.043
1%	Upper	0.298	0.213	0.213	0.213	0.213	0.213	0.213
	Middle	0.043	0.043	0.043	0.043	0.043	0.043	0.043
	Lower	0.553	0.043	0.043	0.043	0.043	0.043	0.043

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

The segment lengths were fixed at 20. The segment parameters considered for the experiments were count local maxima, count of local minima, position of global maxima, position of global minima, value of global minima, value of global minima, number of local minima before global minima, number of local maxima after global minima, number of local minima before global maxima, number of local maxima after global maxima, number of local minima between global minima and global maxima and number of local maxima between global minima and global maxima. The various scenarios for which the experiment was conducted and the REP obtained are given in Table 1. The regions of missing data are indicators of the variations in the lengths of upper and lower sub segments of the segments.

The following observations about the experiment and results are worth mentioning.

- The repletion of the values in the table is a good indicator of the identification of same segment repeatedly irrespective of the increase in number of segments.
- The best results were obtained as the numbers of segments were increased.
- DTW similarity measure is an effective similarity measure for time variant data.
- The increase in segment size also minimizes the error.

V. CONCLUSION

This proposed methodology can be applied to any dataset that suffers from potential loss of data due to instrumental and other unanticipated faults. It cuts out a solution for the problem of missing data by the strategy of Imputation, which involves finding the segment which is most similar to the one with missing data. Proficient feature extraction techniques are adopted to handle the voluminous data which represent continuous time variant data. This minimizes huge computations which need to be done when raw data is used for analysis.SIM performs better than other existing imputation and computational methods because the time complexity is minimized as the use of feature extraction and clustering techniques minimizes the search space to a single cluster. The relative percentage error is also very less which adds to the efficiency of the method. It can perform even better by improvising the clustering algorithms.

REFERENCES

[1] Hui Ding, Goce Trajcevski, Peter Scheurmann, Xiaoyue Wang, Eamonn Keogh: “Querying And Mining Of Time Series Data: Experimental Comparison Of Representations And Distance Measures”, Proceedings of VLDB Endowment, Vol. 1, Pages 1542-1552, 2008

[2] Robert T.Olszewski: “Generalised Feature Extraction for Structural Pattern Recognition in Time Series Data”

[3] Michail Vlachos, Marios Hadjieleftheriou, Dimitrios Gunopulos, Eamonn Keogh: “Indexing Multi-Dimensional Time-Series with Support for Multiple Distance Measures”, Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining”, Pages: 216 – 225,2003

[4] Blaz Strle, Martin Mozina, Ivan Bratko:” Qualitative Approximation To Dynamic Time Warping Similarity Between Time Series Data”, 23rd International Workshop on Qualitative Reasoning 2007

[5] M.Sridharan, N.Gururajan and A.M.S.Ramasamy: “Fuzzy Clustering Analysis to study geomagnetic coastal effects”

[6] M. D. Morse and J. M. Patel. “An efficient and accurate method for evaluating time series similarity”.In SIGMOD Conference, 2007.

[7] I. Popivanov and R. J. Miller. “Similarity Search Over Time-Series Data Using Wavelets”. In ICDE, 2002.

[8] K.-P. Chan and A.-C. Fu.” Efficient Time Series Matching by Wavelets”. ICDE, pages 126–133, 1999

[9] D. Berndt and J. Clifford.” Using Dynamic Time Warping to Find Patterns in Time Series”. In AAAI-94 Workshop on Knowledge Discovery in Databases, pages 359–370, 1994.

[10] B.-K. Yi, H. V. Jagadish, and C. Faloutsos. “Efficient Retrieval of Similar Time Segments Under Time Warping”. In ICDE, pages 201–208, 1998.

AUTHORS PROFILE

F.Sagayaraj Francis is currently working as an Associate Professor in the Department of Computer Science and Engineering, Pondicherry Engineering College. He obtained his M.Tech. degree in Computer Science and Engineering in the year 1997 and Ph.D. degree in Computer Science and Engineering in the year 2008, both from Pondicherry University. His areas of research and interest are Data Management and Information Systems.

B.Vishnupriya has finished her B.Tech in Computer Science and Engineering degree in the current year from Pondicherry University.

Vinolin Deborah Delphin has completed her B.Tech in Computer Science and Engineering degree in the academic year from Pondicherry University.

Saranya Kumari Potluri has done her B.Tech in Computer Science and Engineering degree in the current year from Pondicherry University.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Efficient Node Search in P2P Using Distributed Spanning Tree

P. Victer Paul[*] T. Vengattaraman[*] M.S.Saleem Basha[*] P. Dhavachelvan[*] R. Baskaran[#]
[*]Department of Computer Science, Pondicherry University, Puducherry, India.
[#]Department of Computer Science and Engineering, Anna University, Chennai, India.
{victerpaul, vengat.mailbox, smartsaleem1979, dhavachelvan, baskaran.ramachandran}@gmail.com

Abstract

Peer to Peer (P2P) networks are the important part of the next cohort of Internet, so how to search the node in the P2P networks efficiently is the key problem of the perception of the P2P network. However, the node search process in unstructured P2P is not efficient because the same search message may go through a node multiple times. To ease the complex search and improve the search efficiency, we propose a mechanism using the interconnection structure called Distributed Spanning Tree (DST) which facilitates the P2P Network into a layered structure to improve the node search technique. The performance evaluations of simulation demonstrate that the proposed mechanism can improve the node search efficiency of P2P systems.

Keywords: Peer to Peer, Distributed Spanning Tree, Node Search, Ant colony optimization

1. Introduction

P2P networks are important part of the next generation of Internet, so how to search the resources in the P2P networks efficiently is the key problem of the realization of the P2P network [1]. The advantages of the unstructured P2P networks are that they have lower maintenance overhead and can better adapt to node heterogeneity as well as network dynamics. However, the node search process in unstructured systems is not as efficient because the same search message may go through a node multiple times [2]. For communication in network, operation for search of a particular node is performed often which consumes more message passes even the node identification factor like IP address are known. Since routing information stored by each node is limited that makes the search operation complex in large scale networks. To facilitate the complex search and improve the search efficiency, we propose a novel approach of node search using DST technique. This reduces the message passes required to identify a node and prevent nodes from receiving duplicate search messages and retains the low maintenance overhead for the unstructured system.

From these perspectives, in this paper, it is aimed at developing an optimized node search system using Distributed Spanning Tree (DST), which will reduce the number of message passes required in deterministic environment. Various requirements to reduce the message pass can be achieved by formulating DST in the network. Performance of this improved mechanism is simulated and analyzed using a theme specific environment for Node search. The paper is organized as follows: Section 2 defines the Interconnection Structure used and the respective constrains. Section 3 explains about Formulation of DST in P2P network. Section 4 discusses about proposed mechanisms for Node search in P2P network. Section 5 provides statistical Analysis over result obtained in the Simulation and Section 6 concludes the proposed work with its merit.

2. Interconnection Structure Used

Distributed Spanning Tree (DST) [9, 10] is the interconnection structure we follow to reduce the number of message passes required for node search. DST organizes P2P network into a hierarchy of groups of nodes. The nodes are put together in groups and groups are gathered in higher level recursively. This organization is built on top of routing tables allows the instantaneous

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

creation of spanning trees rooted by many nodes and keeps the load balanced between the nodes [3]. The DST is an overlay structure designed to be scalable [4]. The DST is a tree without bottlenecks which automatically balances the load between its nodes. From [3] and [4], it is possible to achieve economic traffic optimization in Peer Network using DST interconnection structure. So we virtually convert the peer network into distributed spanning trees and each tree should have its root node we call it as *Head Node* (HN) and others are *Leaf Node* (LN). Every LN will hold the details of its own HN. Likewise every HN will hold the complete details regarding its LNs and all other HNs in the network. The details stored in HNs and LNs in the DST is used to enhance the node search operation efficiently with minimum message pass. During the formulation of DST in Peer network LNs and HNs are chosen randomly and dynamically with some requirement criteria which improve the Fault Tolerance of the system.

To enhance the efficiency of DST in Node Search, we focus on a kind of randomized search heuristics, namely Ant Colony Optimization (ACO). ACO [5, 6, 7, 8], is a powerful heuristic approach to solve combinatorial optimization problems such as the TSP, Routing in telecommunication networks. So applying ACO approach can enhance the effective routing of message (at low cost) in the network which in-tern reduces the number of message pass required.

3. Formulation of DST

In a large P2P network, formation of DST is complex. We elucidate the DST formation in the P2P network using five procedures.

Firstly *initialize_DST* is a procedure which initializes DST by creating Head Node (HN) in Peer network based on some test criteria(s). The criteria(s) to be checked can be user approval, traffic on a particular region, etc., and the procedure creates an array on each HN to hold its LN details. If the criteria(s) fail a variable on the node is created. Each HN is provided with unique Priority Number (PN) to provide write priority among the HNs. *initialize_DST is also a* procedure to set its HN id as their own id and then it calls the procedure *probe_DST()*.

The procedure *probe_DST(),* which is called by every HN creates *probe* message and set 'id' field of message as its own id and flood the message to all peers it is connected.

On receiving a message every peer execute the procedure *msg_recieve_DST(msg)* where 'msg' is the received message. During DDST formation it should be possible to get any one of the two types of messages, the *probe message* or *reply message*:

If there is a *probe* message, any one of the following would be occurred:

Case-1: The message is received by a HN: It is just discarded.

Case-2a: The message is received by a LN which is not under any HN: LN stores the Head variable as the id which it read from pmsg. Then call the procedures *probe_reply_DST(N(id))* and *probe_forword_DST(pmsg).*

Case-2b: The message is received by a LN, which is under any HN: It is just discarded.

If there is a *reply* message, any one of the following would be occurred:

Case-1: The message is received by a LN: It just forwards it to the node bearing the id 'id'.

Case-2: If the message is received by a HN: It reads 'dest' from 'rmsg', if 'dest' equals N(id) it shows required HN is reached. It read 'id' from 'rmsg' and add it to its array, otherwise it is forwarded to N(id).

Procedure *probe_reply_DST(N(id))* is called by LN to reply to its HN. The LN creates a *reply* message. The 'id' and 'dest' fields of the *reply* message is set to be, the 'id' of the LN and the 'id' of the HN respectively. After the *reply* message sent to HN, the LN calls the Procedure *probe_forword_DST(pmsg)* to flood the *probe* message to all the peers except the peer from where it was received.

After the completion of these five procedures the Peer Network will be in required DST structure.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Definition 1. Let $n(DST_{msgpass})$ be the Number of Message Pass Required to form DST in the Peer Network and it can be defined as,

$$n(DST_{msgpass}) = ((L/P)*P*M)+((L/P)*N) \quad (1)$$

In equation (1), *'L/P'* gives the number of DST formed in the network which is also equal to *n(HN)*. So equation (1) can be rewritten as,

$$n(DST_{msgpass}) = (L*M)+(n(HN)*N) \quad (2)$$

where,

- 'N' be the number of message pass between one HN and another HN
- 'M' be the number of message pass between HN and LN
- 'L' be the number of Peers in the Network
- 'P' be the number of LN under HN (consider equal number of LN for all HN)
- '(L/P)' be the number of DST in the Peer Network (or equals n(HN))
- It can be interpretable that $1 \le M \le N \le L$ and $1 \le P \le L$

In words, Total number of Message Pass required to form DST in the Peer Network $n(DST_{msgpass})$ is equal to sum of products between number of Peers in the Network and the number of message pass between HN and LN and between the number of HNs in the Network and the number of message pass between one HN and another HN in the Network.

4. Proposed Efficient Node Search in Large scale P2P network using DST

P2P networks are an important part of the next generation of Internet, so how to search the node in the P2P networks efficiently is the key problem of the realization of the P2P network. While studying the search technology of different mechanisms, this paper proposes a method to improve the node search in a large scale P2P network using the systematic DST approach. We propose a Node search algorithm in fig. 1 which consists of five procedures *Request(), toAllHeads(), Found(), Reply() and Receive().*

Procedure *Request()* is invoked by any node in network which want to identify a node for communication. This procedure has two arguments the source node *v* and destination node *d* and creates *doFind* message with source and destination details and sent to HN of source.

Procedure *toallHeads()* is invoked by node *v* (should be a HN) when it receives *Request* message to propagate the message to all other HNs in the network using its HN_ARR array.

Procedure *Found()* is invoked by the HN which hold the requested destination node *d*. This procedure create the *foundDest* message and set fields 'Dest' and 'Head2' as id of destination node d and id of current HN.

Procedure *Reply()* is called by the Requester HN to forward the *foundDest* message to the source node *v*.

Procedure *Receive()* is invoked the node *v*, when it receives any message. The received message should be any of three variants; *doFind message, toallHeads message and foundDest message.*

If the received message by node *v* is a *doFind* message, any one of the following would be occurred:

Case-1: If the node *v* is HN and 'Head' field of the message is 'id' of the node *v*, then it is procedure *toAllHeads()* is invoked.

Case-2: If case-1 fails, then the node *v* forwards the message to the node with id in the 'Head' field of the message.

If the received message by node *v* is a *toallHeads* message, any one of the following would be occurred:

Case-1: If the node *v* is HN and 'Head2' field of the message is 'id' of the node *v*, then node *v* retrieve the 'd' field of message and matches it with LN entry in its LN_ARR array. There may two possible ways,

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Case-1a: If 'd' field in the message matches with LN entry of array LN_ARR in *v* which means that the search node *d* is under the current HN *v*. Then node *v* invoke *Found()* procedure to intimate the requester HN.

Case-1: If the node *v* is HN and 'Head1' field of the message is 'id' of the node *v*, then it is procedure *Reply()* is invoked.

Case-2: If the node *v* is LN and 's' field of the

Request(v,d)
Step 1: create *doFind* message, *dmsg*
Set *Sour* field as v.ID
Set *Dest* field as d.ID
Set *Head* field as v.Head
Step 2: send(*dmsg*,v.Head)

Found(s,d,HN1,HN2)
Step 1: create *foundDest* message, *fmsg*
Set *Sour* field as dmsg.Sour
Set *Dest* field as dmsg.Dest
Set *Head1* field as HN1
Set *Head2* field as v.ID
Step 2: send (*fmsg*,HN1)

toAllHeads(v,dmsg)
Step 1: $\forall$ HN such that HN $\in$ v.HNs repeat step 2 and step 3
Step 2: create *toallHeads* message, *hmsg*
Set *Sour* field as dmsg.Sour
Set *Dest* field as dmsg.Dest
Set *Head1* field as v.ID
Set *Head2* field as HN.ID
Step 3: send (*hmsg*,HN.ID)

Reply(fmsg)
Step 1: send(fmsg,fmsg.Sour)

Receive(msg,v)
Step 1: check if *msg* = *dmsg* goto step 2, if *msg* = *hmsg* goto step 5
if *msg* = *fmsg* goto step 8
Step 2: check if *v* = HN and *v*.ID = dmsg.Head, goto step 3 else step 4
Step 3: **call toAllHeads(v,*pmsg*)**
Step 4: forward(*pmsg*,pmsg.Head)
Step 5: check if *v* = HN and *v*.ID = *hmsg*.Head2, goto step 6 else step 10
Step 6: $\forall$ LN such that LN $\in$ *v*.LNs, repeat step 7
Step 7: if *hmsg*.d = LN.ID, goto step 8 else step 9
Step 8: **call Found(*hmsg*.s,LN.ID, hmsg.Head1, v.ID)**
Step 9: delete *pmsg*
Step 10: forward(*hmsg*,*hmsg*.Head2)
Step 11: check if *v* = HN and v.ID = *fmsg*.Head1, goto step 12
if *v* = LN and *v*.ID = fmsg.s, goto step 13 else step 14
Step 12: **call Reply(*fmsg*)**
Step 13: // communication starts
Step 14: forward(*fmsg*,fmsg.s)

Figure 1. Proposed Node Search algorithm using DST structure

Case-1b: If no matches found in LN_ARR array of HN *v*, then it discards the message.

Case-2: If case-1 fails, then the node *v* forwards the message to the node with id in the 'Head2' field of the message.

If the received message by node *v* is a *foundDest* message, any one of the following would be occurred:

message is 'id' of the node *v*, then it shows that the destination *d* is found. The node *v* retrieves 'Head1', 'Head2' fields of message which it uses for communication with destination *d* through destination HN *Head2*.

Case-3: If both case-1 and case-2 fails, then the node *v* forwards the message to the node with id in the 's' field of the message.

Thus the Proposed Node search algorithm uses the details stored by every HNs regarding its LNs.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

This approach makes the node search operation in a large scale P2P network more efficient and economic with very minimum number of message passes and reduced routing table entry in every node in the network.

Definition 2. Let P_i be a peer in Peer Network. Then the number of routing table entries required for node P_i in the DST Peer network to route any message it receives is $n(P_i(RT\_entry))$ and given as,

$$n(P_i(RT\_entry)) = n(HN) + \{ n(HN_i(LN\_ARR)) \cup r(P_i) \} \quad (3)$$

where,

- ' $n(P_i(RT\_entry))$ ' is the number of routing table entries required for node P_i to route any message it receives.
- ' $n(HN)$ ' is the number of HNs in the Network
- ' $n(HN_i(LN\_ARR))$ ' is the number LNs in the LN_ARR of HN of P_i.
- ' $r(P_i)$ ' is the number of nodes adjacent to P_i.

Thus the number of routing table entries required for node P_i to route any message it receives $n(P_i(RT\_entry))$ is sum of number of HNs in the network and combination (without duplicate) of number of LNs in the LN_ARR of HN of P_i and number of nodes adjacent to node P_i.

5. Simulation and Analysis

This section describes the simulation results obtained during the investigation phases. We used OMNeT++, is an object-oriented modular discrete event network simulator. A Peer Network of 100 peers (comp1, comp2, comp3…comp100) interconnected randomly and spread in some distant geographical location to validate the proposed technique. The proposed Node search technique is implemented in the simulated Peer network and performed various fine grained analyses. . It is assumed that the medium have propagation delay of 10 ms.

In our simulation setup of Peer Network with one hundred systems, $n(DST_{msgpass})$ is nearly equals 391 messages which take nearly 5.83 seconds for the formation of complete DST (i.e.) to organize P2P network into a hierarchy of groups of nodes.

5.1 DST Routing Scheme

In this scheme, to route message from source to destination we used Static Routing technique. From the simulation we gathered different criteria measures and tabulated in Table I which show that the node search operations performed efficiently and economically in DST Peer Network than that of Typical Peer Network. Analyses of large P2P networks (with more than 100 peers) are cautiously derived from the simulated network. On simulation proceeds, it is observed that for the time period of 10 seconds number of node search operations performed in Typical Peer network and DST Peer network are 19 and 32 respectively. Thus the efficiency of node search operation in Peer network can be improved nearly 68.9% by using DST technique.

Table I. Comparison of various criteria measures between Typical and DST Peer Network

S.No	No. of Peers	No. of HNs	Average Routing table entries (approx.)	Typical Peer Network			DST Peer Network			Efficiency
				No. of nodes involved	No. of node search operations	Time taken (sec)	No. of nodes involved	No. of node search operations	Time taken (sec)	
1	10^2	6	24	9	19	10	12	32	10	68.9%
2	10^3	24	67	12	26	5	21	48	5	84.6%
3	10^4	130	202	19	51	2	36	92	2	80.5%
4	10^5	350	601	31	94	1	61	178	1	89.3%

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Table II. Comparison of various criteria measures between DST and ACO optimized DST Peer Network

S.No	No. of Peers	No. of HNs	Average Routing table entries (approx.)	DST Peer Network			ACO DST Peer Network			Efficiency
				No. of nodes involved	No. of node search operations	Time taken (sec)	No. of nodes involved	No. of node search operations	Time taken (sec)	
1	10^2	6	24 + 6	12	32	10	16	44	10	37.5%
2	10^3	24	67 + 23	21	48	5	29	68	5	41.6%
3	10^4	130	202 + 121	36	92	2	48	131	2	42.3%
4	10^5	350	601 + 321	61	178	1	72	214	1	20.2%

5.2 DST with Ant Colony Optimized Routing Scheme

To route message from source to destination Ant Colony Optimized routing technique is followed which improve the message pass efficiency of DST because of dynamically identified optimal route between every HN and LN Peers and alternate optimal route between nodes. Table 2 shows the criteria measure between DST Peer Network and ACO optimized DST Peer Network and from the result analysis, it is evident that the efficiency of node search technique is improved by ACO optimization. On simulation proceeds, it is observed that for the time period of 10 seconds number of node search operations performed in Typical Peer network and DST Peer network are 19 and 32 respectively. Thus the efficiency of node search operation in Peer network can be improved nearly 37.5% by using Ant Colony Optimized DST technique. Analyses of large P2P networks (with more than 100 peers) are cautiously derived from the simulated network.

From the Table I and II, the following observation can be derived.

Observation 1.
Average number of routing table entries required by each peer in DST is more than that of ACO optimized DST.

Proof. In DST network, every peer maintain single optimal path for a destination whereas in ACO technique find more than one optimal path between a single source and destination to increase the reliability on channel failures. In ACO, usually alternate optimal path is maintained between each Peer and all HNs. the number routing table entries required by any peer $n(P_i(RT\_entry))$ given in equation (3) can be expressed as, for DST,

$$n(P_i(RT\_entry)) = n(HN) + \{ \mathrm{n}(\mathrm{HN_i}(LN\_ARR)) \cup \mathrm{r}(\mathrm{P_i}) \} \quad (4)$$

For ACO optimized DST,

$$n(P_i(RT\_entry)) = \{2 * n(HN)\} + \{ \mathrm{n}(\mathrm{HN_i}(LN\_ARR)) \cup \mathrm{r}(\mathrm{P_i}) \} \quad (5)$$

where,

- ' $n(HN)$ ' is the number of HNs in the network
- ' $\mathrm{n}(\mathrm{HN_i}(LN\_ARR))$ ' is the number LNs in the LN_ARR of HN of P_i.
- ' $\mathrm{r}(\mathrm{P_i})$ ' is the number of nodes adjacent to P_i.

From equation (4) and (5), we can conclude that the number routing table entries required by any peer in ACO optimized DST will be more than that of DST.

Observation 2.
With increase in the size of the P2P network, efficiency of DST technique in node search also increases.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

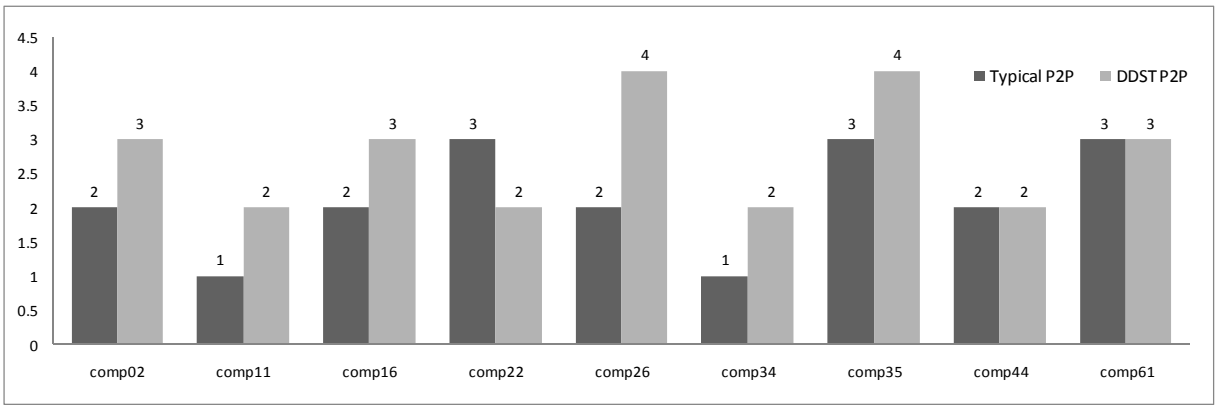


Figure 2. Number node search operations performed by nodes in Typical and DDST P2P

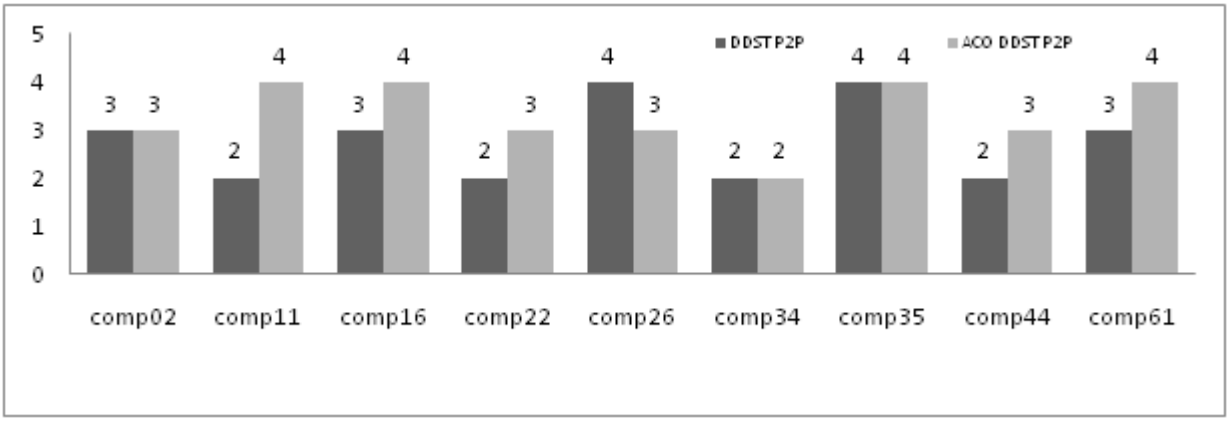


Figure 3 Number node search operations performed by nodes in DDST and ACO DDST P2P

Proof. Number of HNs in the network increases with number of nodes. Consequently the increase in the node search operations is distributed among the HNs. So, large number of node search operations is performed effectively without any bottleneck which increases the overall efficiency of the node search technique using DST. Thus from Table 1, we can conclude that with increase in the size of the P2P network, efficiency of DST technique in node search also increases.

Figure 2 and 3 shows the comparison graph between the number of node search operations performed by nodes in P2P network (only few among the total nodes involved in the search operation). From which we can conclude that the efficiency of node search can be improved using DST technique and ACO technique.

From the simulation analysis, the proposed node search technique optimizes the traffic in efficient manner by reducing the message pass required for search operation in P2P networks.

6. Conclusion

One of the important issues with large P2P networks is that a node search message may go through the same node multiple times, which causes inefficiency of the node search process. We proposed a mechanism to deal with the problem by using interconnection structure DST which will reduce the number of message passes required for node search in the Peer Network. An enhanced algorithm ACO is used to optimize DST, which offered better performance, is also presented. From the simulation analysis, it is shown that in the Peer Network with DST, high performance node search can be achieved than ordinary Peer Network.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Reference

[1]. Mengkun Yang Zongming Fei, Assigning Identifications to Nodes in Unstructured Peer-to-Peer Networks: A Novel Approach to Improving Search Efficiency, Global Telecommunications Conference, 2007.GLOBECOM '07.IEEE.

[2]. Lei Xiao, Liangzhong Shen, ZhongYi Hu, Fei Zhou, "Improving Search in P2P by Location Identification," Environmental Science and Information Application Technology, International Conference on, vol. 1, pp. 519-522, 2009 International Conference on Environmental Science and Information Application Technology, 2009.

[3]. Sylvain Dahan, Jean-Marc Nicod and Laurent Philippe, The Distributed Spanning Tree:A Scalable Interconnection Topology for Efficient and Equitable Traversal, International Symposium on Cluster Computing and the Grid, 2005 IEEE

[4]. S. Dahan, "Distributed Spanning Tree Algorithms for Large Scale Traversals," Proc. 11th Int'l Conf. Parallel and Distributed Systems (ICPADS '05), pp. 453-459, July 2005.

[5]. Runtime Analysis of a Simple Ant Colony Optimization Algorithm Frank Neumann · CarstenWitt, Springer Science+Business Media, LLC 2007

[6]. Dorigo, M., Maniezzo, V., Colorni, A.: The ant system: An autocatalytic optimizing process. Tech.Rep. 91-016 Revised, Politecnico di Milano, Italy (1991)

[7]. M. Dorigo, T. Stützle (2002) The ant colony optimization metaheuristic: Algorithms, applications and advances. In F. Glover, G. Kochenberger (eds) Handbook of Metaheuristics. Kluwer Academic Publishers, Norwell, MA, pp 251-285

[8]. A. Colorni, M. Dorigo, V. Maniezzo (1991) Distributed optimization by ant colonies, In Proceedings of ECAL'91 European Conference on Artificial Life, Elsevier Publishing, Amsterdam, The Netherlands, pp 134-142

[9]. Sylvain Dahan, Jean-Marc Nicod and Laurent Philippe, The Distributed Spanning Tree ,IEEE transactions on parallel and distributed systems, vol. 20, no. 12, december 2009

[10]. A. Rowstron and P. Druschel, "Pastry: Scalable, Decentralized Object Location, and Routing for Large-Scale Peer-to-Peer Systems," Proc. IFIP/ACM Int'l Conf. Distributed Systems Platforms, vol. 2218, pp. 329-350, 2001.

IJCSIS REVIEWERS' LIST

Assist Prof (Dr.) M. Emre Celebi, Louisiana State University in Shreveport, USA

Dr. Lam Hong Lee, Universiti Tunku Abdul Rahman, Malaysia

Dr. Shimon K. Modi, Director of Research BSPA Labs, Purdue University, USA

Dr. Jianguo Ding, Norwegian University of Science and Technology (NTNU), Norway

Assoc. Prof. N. Jaisankar, VIT University, Vellore,Tamilnadu, India

Dr. Amogh Kavimandan, The Mathworks Inc., USA

Dr. Ramasamy Mariappan, Vinayaka Missions University, India

Dr. Yong Li, School of Electronic and Information Engineering, Beijing Jiaotong University, P.R. China

Assist. Prof. Sugam Sharma, NIET, India / Iowa State University, USA

Dr. Jorge A. Ruiz-Vanoye, Universidad Autónoma del Estado de Morelos, Mexico

Dr. Neeraj Kumar, SMVD University, Katra (J&K), India

Dr Genge Bela, "Petru Maior" University of Targu Mures, Romania

Dr. Junjie Peng, Shanghai University, P. R. China

Dr. Ilhem LENGLIZ, HANA Group - CRISTAL Laboratory, Tunisia

Prof. Dr. Durgesh Kumar Mishra, Acropolis Institute of Technology and Research, Indore, MP, India

Jorge L. Hernández-Ardieta, University Carlos III of Madrid, Spain

Prof. Dr.C.Suresh Gnana Dhas, Anna University, India

Mrs Li Fang, Nanyang Technological University, Singapore

Prof. Pijush Biswas, RCC Institute of Information Technology, India

Dr. Siddhivinayak Kulkarni, University of Ballarat, Ballarat, Victoria, Australia

Dr. A. Arul Lawrence, Royal College of Engineering & Technology, India

Mr. Wongyos Keardsri, Chulalongkorn University, Bangkok, Thailand

Mr. Somesh Kumar Dewangan, CSVTU Bhilai (C.G.)/ Dimat Raipur, India

Mr. Hayder N. Jasem, University Putra Malaysia, Malaysia

Mr. A.V.Senthil Kumar, C. M. S. College of Science and Commerce, India

Mr. R. S. Karthik, C. M. S. College of Science and Commerce, India

Mr. P. Vasant, University Technology Petronas, Malaysia

Mr. Wong Kok Seng, Soongsil University, Seoul, South Korea

Mr. Praveen Ranjan Srivastava, BITS PILANI, India

Mr. Kong Sang Kelvin, Leong, The Hong Kong Polytechnic University, Hong Kong

Mr. Mohd Nazri Ismail, Universiti Kuala Lumpur, Malaysia

Dr. Rami J. Matarneh, Al-isra Private University, Amman, Jordan

Dr Ojesanmi Olusegun Ayodeji, Ajayi Crowther University, Oyo, Nigeria

Dr. Riktesh Srivastava, Skyline University, UAE

Dr. Oras F. Baker, UCSI University - Kuala Lumpur, Malaysia

Dr. Ahmed S. Ghiduk, Faculty of Science, Beni-Suef University, Egypt
and Department of Computer science, Taif University, Saudi Arabia

CALL FOR PAPERS
International Journal of Computer Science and Information Security
IJCSIS 2010
ISSN: 1947-5500
http://sites.google.com/site/ijcsis/

International Journal Computer Science and Information Security, now at its sixth edition, is the premier scholarly venue in the areas of computer science and security issues. IJCSIS 2010 will provide a high profile, leading edge platform for researchers and engineers alike to publish state-of-the-art research in the respective fields of information technology and communication security. The journal will feature a diverse mixture of publication articles including core and applied computer science related topics.

Authors are solicited to contribute to the special issue by submitting articles that illustrate research results, projects, surveying works and industrial experiences that describe significant advances in the following areas, but are not limited to. Submissions may span a broad range of topics, e.g.:

Track A: Security

Access control, Anonymity, Audit and audit reduction & Authentication and authorization, Applied cryptography, Cryptanalysis, Digital Signatures, Biometric security, Boundary control devices, Certification and accreditation, Cross-layer design for security, Security & Network Management, Data and system integrity, Database security, Defensive information warfare, Denial of service protection, Intrusion Detection, Anti-malware, Distributed systems security, Electronic commerce, E-mail security, Spam, Phishing, E-mail fraud, Virus, worms, Trojan Protection, Grid security, Information hiding and watermarking & Information survivability, Insider threat protection, Integrity
Intellectual property protection, Internet/Intranet Security, Key management and key recovery, Language-based security, Mobile and wireless security, Mobile, Ad Hoc and Sensor Network Security, Monitoring and surveillance, Multimedia security ,Operating system security, Peer-to-peer security, Performance Evaluations of Protocols & Security Application, Privacy and data protection, Product evaluation criteria and compliance, Risk evaluation and security certification, Risk/vulnerability assessment, Security & Network Management, Security Models & protocols, Security threats & countermeasures (DDoS, MiM, Session Hijacking, Replay attack etc,), Trusted computing, Ubiquitous Computing Security, Virtualization security, VoIP security, Web 2.0 security, Submission Procedures, Active Defense Systems, Adaptive Defense Systems, Benchmark, Analysis and Evaluation of Security Systems, Distributed Access Control and Trust Management, Distributed Attack Systems and Mechanisms, Distributed Intrusion Detection/Prevention Systems, Denial-of-Service Attacks and Countermeasures, High Performance Security Systems, Identity Management and Authentication, Implementation, Deployment and Management of Security Systems, Intelligent Defense Systems, Internet and Network Forensics, Large-scale Attacks and Defense, RFID Security and Privacy, Security Architectures in Distributed Network Systems, Security for Critical Infrastructures, Security for P2P systems and Grid Systems, Security in E-Commerce, Security and Privacy in Wireless Networks, Secure Mobile Agents and Mobile Code, Security Protocols, Security Simulation and Tools, Security Theory and Tools, Standards and Assurance Methods, Trusted Computing, Viruses, Worms, and Other Malicious Code, World Wide Web Security, Novel and emerging secure architecture, Study of attack strategies, attack modeling, Case studies and analysis of actual attacks, Continuity of Operations during an attack, Key management, Trust management, Intrusion detection techniques, Intrusion response, alarm management, and correlation analysis, Study of tradeoffs between security and system performance, Intrusion tolerance systems, Secure protocols, Security in wireless networks (e.g. mesh networks, sensor networks, etc.), Cryptography and Secure Communications, Computer Forensics, Recovery and Healing, Security Visualization, Formal Methods in Security, Principles for Designing a Secure Computing System, Autonomic Security, Internet Security, Security in Health Care Systems, Security Solutions Using Reconfigurable Computing, Adaptive and Intelligent Defense Systems, Authentication and Access control, Denial of service attacks and countermeasures, Identity, Route and

Location Anonymity schemes, Intrusion detection and prevention techniques, Cryptography, encryption algorithms and Key management schemes, Secure routing schemes, Secure neighbor discovery and localization, Trust establishment and maintenance, Confidentiality and data integrity, Security architectures, deployments and solutions, Emerging threats to cloud-based services, Security model for new services, Cloud-aware web service security, Information hiding in Cloud Computing, Securing distributed data storage in cloud, Security, privacy and trust in mobile computing systems and applications, **Middleware security & Security features:** middleware software is an asset on
its own and has to be protected, interaction between security-specific and other middleware features, e.g., context-awareness, **Middleware-level security monitoring and measurement:** metrics and mechanisms for quantification and evaluation of security enforced by the middleware, **Security co-design:** trade-off and co-design between application-based and middleware-based security, **Policy-based management:** innovative support for policy-based definition and enforcement of security concerns, **Identification and authentication mechanisms:** Means to capture application specific constraints in defining and enforcing access control rules, **Middleware-oriented security patterns:** identification of patterns for sound, reusable security, **Security in aspect-based middleware:** mechanisms for isolating and enforcing security aspects, **Security in agent-based platforms:** protection for mobile code and platforms, Smart Devices: Biometrics, National ID cards, Embedded Systems Security and TPMs, RFID Systems Security, Smart Card Security, Pervasive Systems: Digital Rights Management (DRM) in pervasive environments, Intrusion Detection and Information Filtering, Localization Systems Security (Tracking of People and Goods), Mobile Commerce Security, Privacy Enhancing Technologies, Security Protocols (for Identification and Authentication, Confidentiality and Privacy, and Integrity), Ubiquitous Networks: Ad Hoc Networks Security, Delay-Tolerant Network Security, Domestic Network Security, Peer-to-Peer Networks Security, Security Issues in Mobile and Ubiquitous Networks, Security of GSM/GPRS/UMTS Systems, Sensor Networks Security, Vehicular Network Security, Wireless Communication Security: Bluetooth, NFC, WiFi, WiMAX, WiMedia, others

This Track will emphasize the design, implementation, management and applications of computer communications, networks and services. Topics of mostly theoretical nature are also welcome, provided there is clear practical potential in applying the results of such work.

Track B: Computer Science

Broadband wireless technologies: LTE, WiMAX, WiRAN, HSDPA, HSUPA, Resource allocation and interference management, Quality of service and scheduling methods, Capacity planning and dimensioning, Cross-layer design and Physical layer based issue, Interworking architecture and interoperability, Relay assisted and cooperative communications, Location and provisioning and mobility management, Call admission and flow/congestion control, Performance optimization, Channel capacity modeling and analysis, Middleware Issues: Event-based, publish/subscribe, and message-oriented middleware, Reconfigurable, adaptable, and reflective middleware approaches, Middleware solutions for reliability, fault tolerance, and quality-of-service, Scalability of middleware, Context-aware middleware, Autonomic and self-managing middleware, Evaluation techniques for middleware solutions, Formal methods and tools for designing, verifying, and evaluating, middleware, Software engineering techniques for middleware, Service oriented middleware, Agent-based middleware, Security middleware, Network Applications: Network-based automation, Cloud applications, Ubiquitous and pervasive applications, Collaborative applications, RFID and sensor network applications, Mobile applications, Smart home applications, Infrastructure monitoring and control applications, Remote health monitoring, GPS and location-based applications, Networked vehicles applications, Alert applications, Embeded Computer System, Advanced Control Systems, and Intelligent Control : Advanced control and measurement, computer and microprocessor-based control, signal processing, estimation and identification techniques, application specific IC's, nonlinear and adaptive control, optimal and robot control, intelligent control, evolutionary computing, and intelligent systems, instrumentation subject to critical conditions, automotive, marine and aero-space control and all other control applications, Intelligent Control System, Wiring/Wireless Sensor, Signal Control System. Sensors, Actuators and Systems Integration : Intelligent sensors and actuators, multisensor fusion, sensor array and multi-channel processing, micro/nano technology, microsensors and microactuators, instrumentation electronics, MEMS and system integration, wireless sensor, Network Sensor, Hybrid

Sensor, Distributed Sensor Networks. Signal and Image Processing : Digital signal processing theory, methods, DSP implementation, speech processing, image and multidimensional signal processing, Image analysis and processing, Image and Multimedia applications, Real-time multimedia signal processing, Computer vision, Emerging signal processing areas, Remote Sensing, Signal processing in education. Industrial Informatics: Industrial applications of neural networks, fuzzy algorithms, Neuro-Fuzzy application, bioInformatics, real-time computer control, real-time information systems, human-machine interfaces, CAD/CAM/CAT/CIM, virtual reality, industrial communications, flexible manufacturing systems, industrial automated process, Data Storage Management, Harddisk control, Supply Chain Management, Logistics applications, Power plant automation, Drives automation. Information Technology, Management of Information System : Management information systems, Information Management, Nursing information management, Information System, Information Technology and their application, Data retrieval, Data Base Management, Decision analysis methods, Information processing, Operations research, E-Business, E-Commerce, E-Government, Computer Business, Security and risk management, Medical imaging, Biotechnology, Bio-Medicine, Computer-based information systems in health care, Changing Access to Patient Information, Healthcare Management Information Technology. Communication/Computer Network, Transportation Application : On-board diagnostics, Active safety systems, Communication systems, Wireless technology, Communication application, Navigation and Guidance, Vision-based applications, Speech interface, Sensor fusion, Networking theory and technologies, Transportation information, Autonomous vehicle, Vehicle application of affective computing, Advance Computing technology and their application : Broadband and intelligent networks, Data Mining, Data fusion, Computational intelligence, Information and data security, Information indexing and retrieval, Information processing, Information systems and applications, Internet applications and performances, Knowledge based systems, Knowledge management, Software Engineering, Decision making, Mobile networks and services, Network management and services, Neural Network, Fuzzy logics, Neuro-Fuzzy, Expert approaches, Innovation Technology and Management : Innovation and product development, Emerging advances in business and its applications, Creativity in Internet management and retailing, B2B and B2C management, Electronic transceiver device for Retail Marketing Industries, Facilities planning and management, Innovative pervasive computing applications, Programming paradigms for pervasive systems, Software evolution and maintenance in pervasive systems, Middleware services and agent technologies, Adaptive, autonomic and context-aware computing, Mobile/Wireless computing systems and services in pervasive computing, Energy-efficient and green pervasive computing, Communication architectures for pervasive computing, Ad hoc networks for pervasive communications, Pervasive opportunistic communications and applications, Enabling technologies for pervasive systems (e.g., wireless BAN, PAN), Positioning and tracking technologies, Sensors and RFID in pervasive systems, Multimodal sensing and context for pervasive applications, Pervasive sensing, perception and semantic interpretation, Smart devices and intelligent environments, Trust, security and privacy issues in pervasive systems, User interfaces and interaction models, Virtual immersive communications, Wearable computers, Standards and interfaces for pervasive computing environments, Social and economic models for pervasive systems, Active and Programmable Networks, Ad Hoc & Sensor Network, Congestion and/or Flow Control, Content Distribution, Grid Networking, High-speed Network Architectures, Internet Services and Applications, Optical Networks, Mobile and Wireless Networks, Network Modeling and Simulation, Multicast, Multimedia Communications, Network Control and Management, Network Protocols, Network Performance, Network Measurement, Peer to Peer and Overlay Networks, Quality of Service and Quality of Experience, Ubiquitous Networks, Crosscutting Themes – Internet Technologies, Infrastructure, Services and Applications; Open Source Tools, Open Models and Architectures; Security, Privacy and Trust; Navigation Systems, Location Based Services; Social Networks and Online Communities; ICT Convergence, Digital Economy and Digital Divide, Neural Networks, Pattern Recognition, Computer Vision, Advanced Computing Architectures and New Programming Models, Visualization and Virtual Reality as Applied to Computational Science, Computer Architecture and Embedded Systems, Technology in Education, Theoretical Computer Science, Computing Ethics, Computing Practices & Applications

Authors are invited to submit papers through e-mail ijcsiseditor@gmail.com. Submissions must be original and should not have been published previously or be under consideration for publication while being evaluated by IJCSIS. Before submission authors should carefully read over the journal's Author Guidelines, which are located at http://sites.google.com/site/ijcsis/authors-notes .

ISSN 1947 5500